

/2024



</annual>

{report}





Table of Contents

2 Foreword	5 Collaborating with CERN openlab	30 Emerging Technologies
3 About CERN openlab	6 Our members	32 Evaluation of Cerabyte: Archival Data Storage Technology using Ceramic Nanolayers
4 Our people	8 R&D directions	34 interTwin: Co-designing and Prototyping an Interdisciplinary Digital Twin Engine
10 Sustainable Infrastructures		36 BioDynaMo: Biology Dynamics Modeller
12 Integration of Oracle Cloud Resources into CERN IT Business Continuity & Disaster Recovery		38 EMP2: Environmental Modelling and Prediction Platform
14 Facilitate and Automatize Kubernetes Operations		39 Starting in 2025
16 Real-time Data Processing for Level-1 Trigger Scouting at CMS using CXL Memory-Lake Architecture		40 Next-Generation Exascale Flash Storage
18 Next Generation Archiver for WinCC OA		41 Oracle Kubernetes Operator
20 Data Analytics for Industrial Control Systems		42 Cost Optimization and Sustainability for Public Cloud Provider
22 Heterogeneous Architectures Testbed		43 Anomaly Detection for Ultra Low Latency Event Selection at the LHC
24 SPECTRUM		44 Digital Twin: Data Science Engine
26 Center of Excellence on AI and Simulation-Based Engineering at Exascale (CoE RAISE)		45 Applied Multi-Disciplinary AI on High-Performance Computing
28 Quantum Databases for Dynamic Data Storage		46 ODISSEE: Online Data Intensive Solutions for Science in the Exabytes Era
	48 Strategic Partnerships Incubator	56 Communication & Outreach
	52 Training & Education	58 Publications & Presentations
	54 Technical Workshop	

Foreword

It is with great enthusiasm that I present the CERN openlab 2024 Annual Report, a reflection of the remarkable achievements and collaborative spirit that define our unique collaboration. This year, we have once again demonstrated the transformative power of bringing together industry, researchers, and academic institutions to push the boundaries of scientific and technological innovation.



As CERN continues its mission to unravel the mysteries of the universe, CERN openlab serves as a vital bridge, ensuring that state-of-the-art computing technologies are harnessed to meet the immense challenges of data analysis and simulation at unprecedented scales. In 2024, our partnerships deepened across domains such as artificial intelligence, high-performance computing, digital twins, and advanced storage solutions, enabling breakthroughs that not only empower fundamental physics research but also drive broader societal impact.

At the heart of CERN openlab's success lies its collaborative essence. By fostering an environment where expertise is shared freely and challenges are tackled collectively, we have created a thriving ecosystem where ideas are transformed into tangible advancements. This year's projects highlight the immense potential of such partnerships, from improving data-processing pipelines to developing innovative applications of emerging technologies that extend beyond high-energy physics.

Maria Girone – Head of CERN openlab



"CERN openlab is key to CERN's mission, serving as a bridge between cutting-edge industry technologies and the computing needs of our scientific community. By fostering innovation and collaboration, it ensures we have the tools to process and analyse the immense volumes of data generated by our experiments, driving discoveries that expand our understanding of the universe."

Joachim Mnich – CERN Director for Research and Computing

"CERN openlab is an essential pillar of the CERN IT department, driving innovation and enabling us to tackle the unprecedented computing challenges of modern physics. Through collaboration with industry and academia, it brings frontier technologies into our infrastructure, ensuring we remain agile, efficient, and ready to support the scientific discoveries of tomorrow."

Enrica Porcari – Head of CERN IT Department



About CERN openlab

CERN openlab is a unique public-private partnership that has been at the forefront of technological innovation since its establishment in 2001. Designed as a bridge between CERN and leading companies in the IT and computing sectors, CERN openlab provides a collaborative platform to tackle the unprecedented computing challenges associated with the organization's scientific mission. By working hand-in-hand with industry and academic partners, CERN openlab enables the testing and optimization of cutting-edge technologies in a demanding, research-driven environment.

Over the years, CERN openlab has evolved into an innovative collaboration, playing a role in addressing the computing requirements of CERN's experiments, being instrumental in supporting scientific breakthroughs. Its collaborative projects have explored areas such as high-performance computing, artificial intelligence, quantum technologies, and scientific digital twins, ensuring CERN stays at the forefront of computational science.

CERN openlab operates within structured three-year phase cycles designed to systematically assess technological evolution, anticipate future needs, and delineate overarching thematic priorities. This approach ensures the maintenance of a relevant and current research programme, fostering effective collaborations and innovative advancements.

Upon joining CERN openlab, members gain access to a unique ecosystem characterised by unparalleled computing challenges, ground-breaking scientific endeavours, and pioneering minds. This environment is essential for the development and demonstration of emerging technologies, providing a platform for industry leaders to showcase their potential and validate solutions through realistic, demanding use cases. This process often leads to tangible enhancements in product features and capabilities.

CERN openlab's enduring success lies in its ability to unite worldwide leading experts and frontier technologies to tackle complex problems at the intersection of science and technology. As it looks ahead to the next years, it remains a beacon of collaboration, innovation, and excellence, ensuring that CERN remains a global leader in both scientific research and technological advancement.



Our people



Maria Girone
Head of CERN openlab



Luca Atzori
CTO for Computing



Antonio Nappi
CTO for Platforms and Workflows



Luca Mascetti
CTO for Storage



Thomas Owen James
CTO for AI and Edge Devices



Eric Wulff
Deputy CTO



Mariana Velho
Communication Manager



Killian Verder
CTO Office Administrative



Valentina Clavel
Finance Manager

Collaborating with CERN openlab



"With our data processing R&D we benefit significantly from CERN openlab through its connections to relevant industry partners for joint R&D, through access to future and innovative hardware and HPC resources and expertise. We also rely on CERN openlab's engagement to 'keep things running' by providing technical and administrative support, supplying us with bright students and providing us with smoothly operating monster hardware."

Axel Naumann – Senior Applied Physicist at CERN EP-SFT



"Having recently joined CERN openlab, as project leader of the ATLAS Trigger and Data Acquisition (TDAQ) system, I really appreciate the boost CERN openlab provides by enabling the 92 institutes involved in TDAQ to interact effectively with industrial partners. These contacts can play an important role in our upgrade programme for the High-Luminosity LHC. The CERN openlab team is very committed to helping us, and in a few months we have managed to start interesting projects in the field of AI, state-of-the-art GPU architectures and the deployment of anomaly detection algorithms in the Level-1 Trigger system, which is currently taking data. At the end of 2025, ATLAS will have to make important architectural choices, and the CERN openlab team allows us to have a much broader view of what is at the cutting edge of technology."

Stefano Veneziano – Senior Applied Physicist at CERN EP-ADT-TR



"CERN openlab collaboration provides a great framework for developing solutions to some of CERN's unique computing challenges while leveraging the most recent cutting-edge technologies of our partners. The strength of openlab lies in innovation and their highly driven and engaged teams who understand the value of bridging academia and industry. It is an excellent opportunity for young professionals to work at CERN and to contribute to innovative projects with industry leaders."

Aleksandra Wardzinska – CERN IT Platforms & Workflows Group Leader



"CERN openlab is a very special partner for IDEAS4HPC. Our mission is to facilitate the access of scientists from under-represented groups to scientific and technical training of the highest-ranking institutions. We are very pleased to be able to support female students to attend the renowned CERN openlab Summer Students Programme."

Marie-Christine Sawley – HPC and Exascale Advisor at ICES Foundation Geneva



"By engaging in collaboration with leading ICT companies through CERN openlab, the IT department and groups are able to integrate complementary knowledge, skills and resources. It is a fantastic opportunity to work together to create new ideas and develop even more powerful technologies that will advance both science and industry."

Eva Dafonte Pérez – CERN IT Databases and Analytics Group Leader



"CERN openlab is an essential platform for implementing the R&D roadmap for storage and data management for the High-Luminosity LHC era, and beyond. Our ongoing industrial collaborations within CERN openlab include understanding transformational potential of new media for long-term data archival, exploration of storage requirements for ML/AI applications and optimization of high-performance, ultra-dense solid-state storage for physics data workflows. These projects are a vital part of our strategy to ensure that our storage capabilities continue to meet the requirements of CERN's scientific programme, now and in the future."

Jakub T. Mościcki – CERN IT Storage & Data Management Group Leader

Our members

Members gain access to a unique ecosystem characterized by unparalleled computing challenges, ground-breaking scientific endeavours, and pioneering minds. Providing a platform for industry leaders to showcase their potential and validate solutions through realistic, demanding use cases. This process often leads to tangible enhancements in product features and capabilities.

ORACLE

micron

PURE STORAGE

SIEMENS

cerabyte

Johnson & Johnson

E4
COMPUTER
ENGINEERING

intel

SIM NS
FOUNDATION

Pasqual

INFN

 **UNIVERSITÀ
DEGLI STUDI DI TRIESTE**

R&D directions

Throughout its history, CERN has been at the forefront of big data scientific research, with CERN openlab playing a pivotal role in tackling the associated computing challenges. By fostering collaborations with industry and research organisations, CERN openlab empowers the HEP community in its research endeavours. In response to the evolving landscape of scientific research, including the advent of exascale computing, CERN openlab spearheads efforts to enhance and scale up IT infrastructure to tackle the upcoming data challenges.

Sustainable Infrastructures

Heterogeneous Computing, Platforms and HPC systems

Computing Architectures and Software Engineering

Advanced Storage, Data Management and Networks

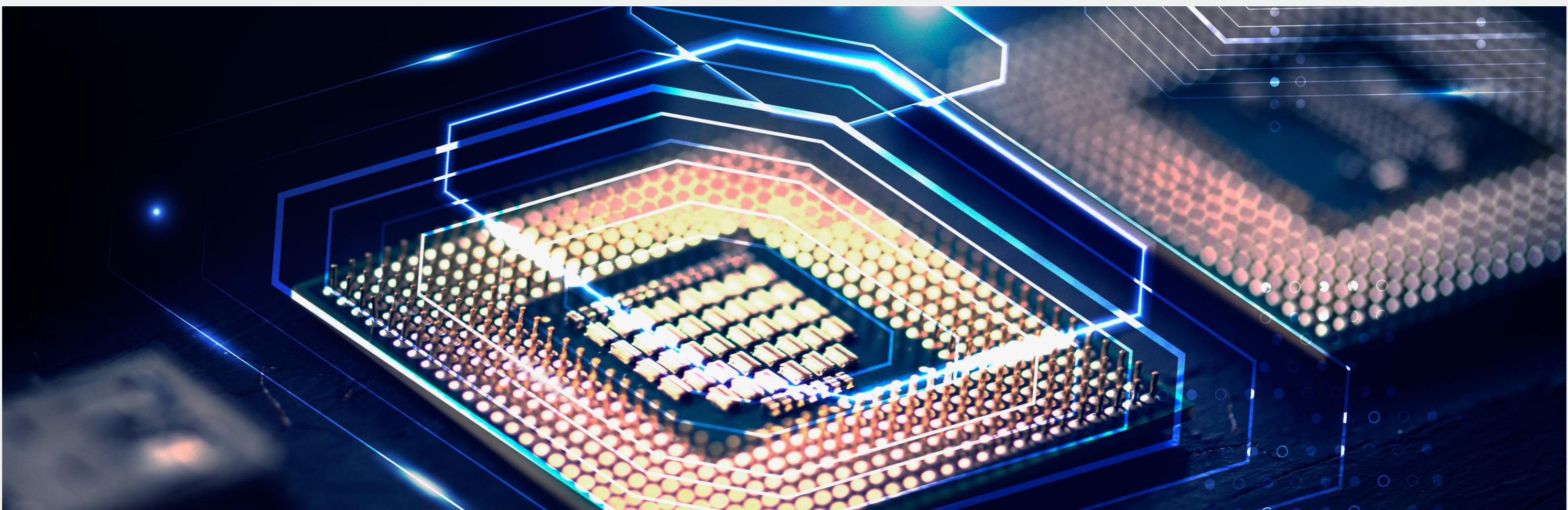
Infrastructures and Techniques for Artificial Intelligence

Applications for Society and Environment

Emerging Technologies

Digital Twins

New Materials for Long-Term Digital Storage



Sustainable Infrastructures



Integration of Oracle Cloud Resources into CERN IT Business Continuity & Disaster Recovery

Aims to solidify a robust disaster recovery plan for CERN's critical on-premises Oracle databases, leveraging advancements made in infrastructure, security, and automation. The project aims to ensure seamless database fail-over and recovery through Oracle Data Guard Database replicas operating asynchronously in Oracle Cloud Infrastructure (OCI). Enhanced configurations, including Terraform automation and multi-region networking capabilities, support this effort while maintaining full data governance under CERN's control.



Overview

CERN's critical databases are foundational to its mission, requiring stringent protection against prolonged infrastructure outages. Significant progress has already been achieved in replicating data securely to off-site OCI regions using private links via the GÉANT network and Oracle Fast Connect. Improving automation workflows has laid the groundwork for a resilient disaster recovery strategy. These efforts ensure not only data preservation but also operational flexibility for reconstruction and fail-over scenarios. The focus now shifts to optimizing costs, refining cloud integration, and finalizing the framework for production-level reliability.

Highlights in 2024

A production-level switch-over for selected critical Oracle databases was planned for 2025Q1 during year-end technical stop and key metrics will be captured to ensure precise expectations for reconstruction and fail-over scenarios.

A complete tear-down and rebuild of off-site OCI resources was initiated several times to validate configurations and ensure comprehensive documentation for future use.

Terraform automation via OCI stacks is being incorporated into the deployment process to streamline and standardize the tear-down and rebuild of OCI resources, reducing manual intervention and improving efficiency.

Review of data governance, access controls, and life-cycle management processes were conducted to align with CERN's data protection policies.

CERN expanded its OCI tenancy networking into additional regions to enable provisioning of GPU-based resources. This effort supports advanced computational needs beyond disaster recovery project and enhances resource flexibility across regions.

Next Steps

FinOps practices will be implemented to ensure cost optimization and transparency in cloud resource usage. This includes monitoring and managing cloud expenses, analysing cost-performance trade-offs, and identifying opportunities for savings without compromising performance or security.

A comprehensive comparison between cloud-based and on-premise disaster recovery solutions will be conducted. This analysis will focus on factors such as cost efficiency, scalability, performance, security, and long-term sustainability. The insights from this comparison will inform strategic decisions about future infrastructure investments and operations.

In collaboration with the CERN Computer Security Team and the Data Protection Office, security hardening efforts and data privacy review are planned to bolster protection against threats.

Project Coordinator

Miroslav Potocky

Technical Team

Miroslav Potocky
Alexandros Stoumpis

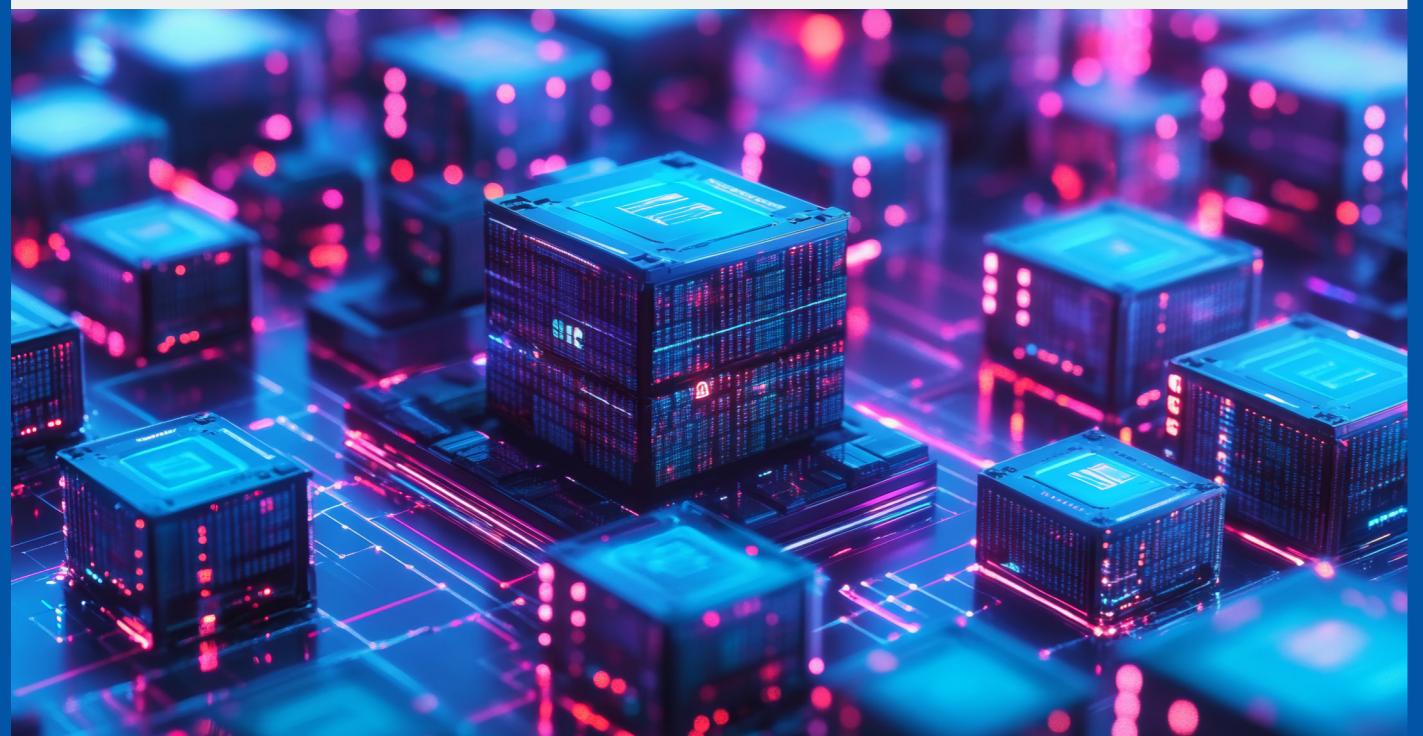
Collaboration Liaisons

Şengül Chardonnereau
Sébastien Hurel
Stefan Jung
John Lathouwers
Cristobal Pedregal-Martin
Eva Dafonte Perez

ORACLE

Facilitate and Automatize Kubernetes Operations

This project aims to develop a tool for statically validating Kubernetes workloads across different Kubernetes versions. This will simplify and automate cluster upgrades by predicting whether a workload is compatible with a new Kubernetes version, mitigating the risks of in-place upgrades and the costs of parallel infrastructure.



Overview

The main issue during upgrade to a newer version of Kubernetes cluster is that we cannot statically determine if our current Kubernetes workloads are going to break in the newer version because of API changes/ deprecations. The only way to determine it is to run it against a new Kubernetes cluster. When you have thousands of pods/ resources this way to perform operations doesn't scale. The testing should not follow an empiric strategy but take advantage of a static analysis. This is the first step of a wider idea for statistical analysis of service mesh dependencies.

In supporting our infrastructure engineers, as well as engineers worldwide working with Kubernetes, we focused on making workload migration between cluster versions less error-prone, faster, more efficient, and less susceptible to human oversight.

Next Steps

Next year, our focus will be on Phase VIII of openlab. This phase will involve further modernization of the CERN ORDS service, leveraging Oracle's new Kubernetes support. We will also utilize the Oracle Kubernetes Operator to modernize Oracle Database management, simplifying database instance provisioning for developers.

Project Coordinator

Antonio Nappi

Technical Team

Antonio Nappi
Adrian Karasiński
Artur Wiecek

Collaboration Liaisons

Eric Grancher
Cristobal Pedregal-Martin
Miroslav Potocky
Garret Swart
Aleksandra Wardzinska
John Lathouwers

ORACLE

Highlights in 2024

The project related to static analysis of context (kubernetes-diff) is considered completed. A proof of concept was produced that unfortunately couldn't cover the entire domain of possible cases due to limitations of Kubernetes schemas, which are often incomplete. This inability to achieve 100% case coverage makes it impossible to fully realize the project's objectives.

However, we worked to raise awareness within the community by contributing with updates to the Kubernetes project documentation. These changes aim to enhance community understanding, assist feature researchers, and explicitly address this limitation in the appropriate context. This is particularly important since such knowledge is typically familiar only to those directly working on the Kubernetes codebase, while average users may lack this insight.

We collaborated with the Kubernetes API Machinery Special Interest Group (sig-api-machinery) on our contribution [1] to the official Kubernetes documentation, consulting them on optimal dissemination strategies to prevent others from encountering the same issues.

Since active development on the Kubernetes tool was stopped, in agreement with Oracle we decided to work in parallel on other projects of interest for both parties.

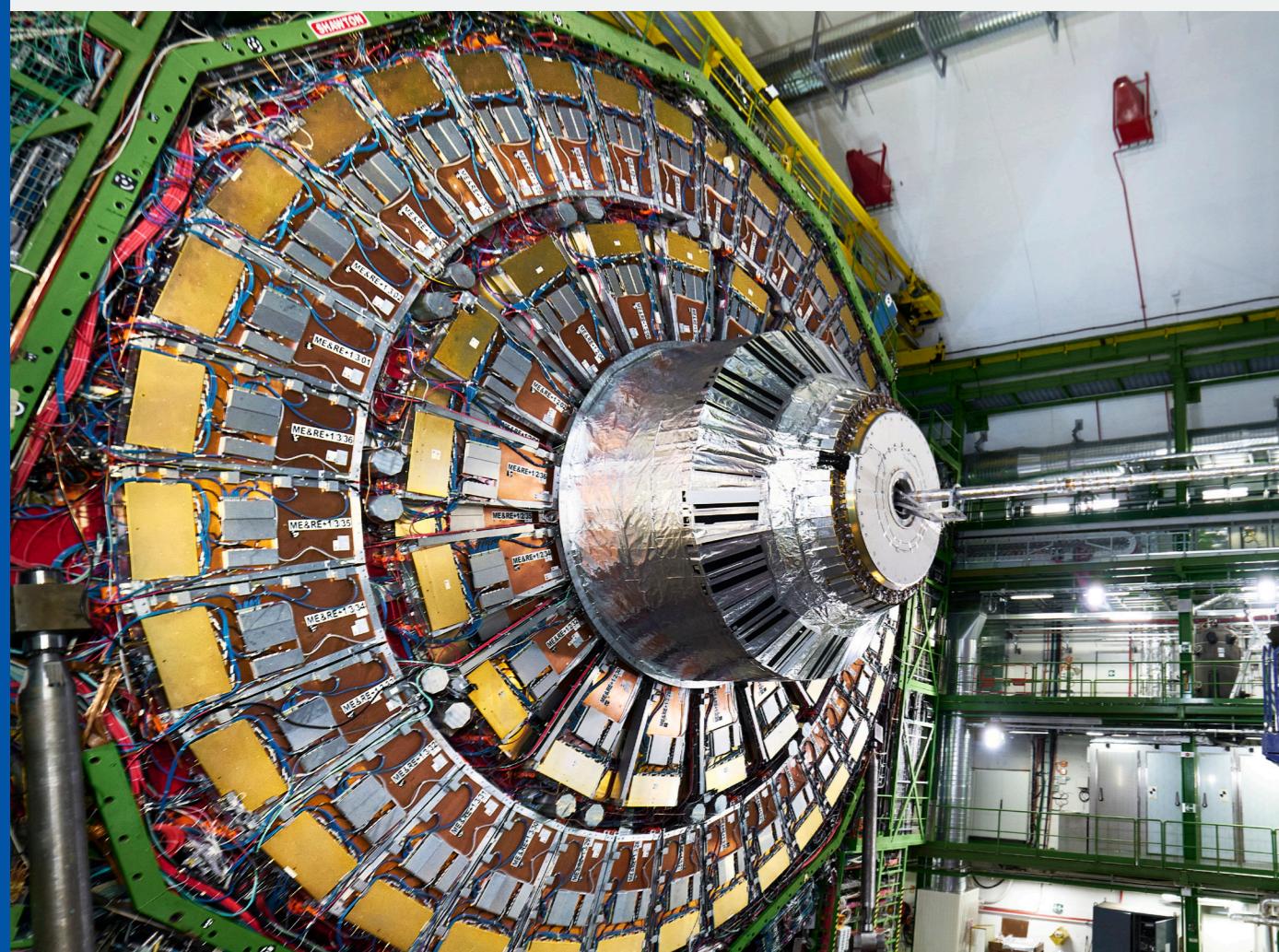
We spent the first half of the year testing and evaluating the Oracle Database Multilingual Engine (Oracle MLE). Our results showed significant performance improvements—several times faster in some cases—for specific use cases. This suggests developers can benefit from using modern languages like JavaScript for server-side procedural logic.

During the second half of the year, we collaborated with Oracle to enhance Oracle REST Data Services (ORDS) authentication, enabling OpenID Connect support for CERN's ORDS service and eliminating the need for custom authentication components. This project remains in progress, with further development planned for the near future.

[1] <https://github.com/kubernetes/website/pull/49025>

Real-time Data Processing for Level-1 Trigger Scouting at CMS using CXL Memory-Lake Architecture

Aims to use the Micron CXL-enabled memory devices as part of the ingestion and data processing chain for the L1 Scouting system at CMS, providing a coherent and seamless access to buffered data from multiple processors and compute accelerators, and a low-latency access/short term storage space for both raw and processed data at scale.



Overview

The Compute Express Link (CXL) protocol is a new alternative protocol that can run over the standard PCIe physical layer, and dynamically multiplexes IO, cache and memory protocols. It is designed to empower a new generation of heterogeneous and disaggregated computing with efficient resource sharing, shared memory pools, enhanced movement of operands and results between accelerators and target devices, and significant latency reduction. CMS intends to profit from the capabilities of this new technology in the online processing solution for the L1 scouting data, and in doing so will pave the way for its utilization in the wider community.

Highlights in 2024

2024 began with extensive benchmarking that proved that the Micron CXL-enabled memory expansion modules, CZ120s with 256 GB capacity, were capable of previously unprecedented sustained memory bandwidth per channel at latencies approaching that of pure DRAM. Following these results, CXL support was added to the data acquisition software of the CMS L1 scouting, allowing data to be written to the memory lake prototype that hosts the Micron devices. In order to improve the capabilities of this system beyond pure direct-attached-memory addressing, we tested and deployed a new open-source CXL-enabled filesystem known as Fabric-attached memory filesystem (FAMFS), developed in tandem with Micron engineers for shared, disaggregated file management.

Three hackathons were hosted throughout the year at the CERN IdeaSquare, where work on the CXL implementation of the data acquisition software was accelerated and spotlighted.

In November, two researchers on the project were able to attend SuperComputing conference in Atlanta, Georgia, USA. In addition to using the opportunity to connect with global leaders in high performance computing, and engage with the forefront of computing innovations, the researchers were able to visit the Micron offices in Atlanta for a hands-on training of their new xCiter near-memory compute FPGA-based platform based on CXL 1.1.

Next Steps

In early 2025, we will receive a full CXL memory-lake system with up to 5.5 TB of CXL-enabled storage (22 modules) in a single chassis, accessed over the first available CXL switch. This memory chassis will be connected to at least four host PCs, which will receive the physics data from the L1 scouting system on multiple 100Gb/s TCP/IP connections.

Project Coordinator

Emilio Meschi

Technical Team

Davide Cristoforetti
Thomas Owen James
Emilio Meschi
Giovanna Lazzari Miotto
Guilherme Paulino

Collaboration Liaisons

Jason Adlard
John Benavides
Robert Bredehoft
Tony Brewer
Emanuele Confalonieri
Henri Courtier
Glen Edwards
John Groves
Andrey Kudryavtsev
Skyler Windh



Next Generation Archiver for WinCC OA

Aims to improve the performance and enhance the archiving functionality of WinCC OA – a SCADA tool used in more than 800 mission-critical control systems at CERN. In addition to supporting new database technologies, it facilitates the use of data analytics techniques on the recorded data, in order to detect anomalies and systematic issues that may impact system operation and maintenance.



Overview

The HL-LHC programme aims to increase the integrated luminosity – and hence the rate of particle collisions – by a factor of ten beyond the LHC’s design value. Monitoring and control systems will therefore become increasingly complex, with unprecedented data throughputs.

Consequently, it is vital to further improve the performance of these systems, and to make use of data analytics algorithms to detect anomalies and anticipate future behaviour. Achieving this involves a number of related lines of work. This project focuses on the development of a modular and future-proof archiving system (NextGen Archiver – NGA) that supports different SQL and NOSQL database technologies to improve the performance and enable new use cases, including data analytics.

Highlights in 2024

In 2024, the team focused on two main objectives: supporting new production deployments of the NGA and developing the TimescaleDB backend, culminating in its first preview release for CERN users.

The deployment of the NGA across more than 500 production WinCC OA systems in the four major LHC experiments – ATLAS, CMS, LHCb, and ALICE – was successfully completed in the first quarter of 2024. Given the number and complexity of the systems involved, significant effort was dedicated to supporting users during and after the migration.

On the development side, the team focused on implementing and testing various functionalities of the TimescaleDB backend and preparing its preview release for deployment in pilot systems. Multiple performance tests were conducted throughout the year using a dedicated TimescaleDB instance provided by the CERN Database On Demand team, demonstrating that significant performance improvements can be achieved through the use of compression and continuous aggregates. Additionally, work on extending the query functionality of WinCC OA, including the Trend widget, is well advanced, with the first improvements already incorporated into the product.

Next Steps

The work on the TimescaleDB backend will continue in 2025, with the goal of releasing a production-grade version in 2026. Particular attention will be given to addressing different use cases for continuous aggregates. The team will also provide support to pilot users in the experiments and gather their feedback. Additionally, various solutions for TimescaleDB high availability and load balancing will be evaluated at a scale that simulates production systems.

Project Coordinator

Rafal Kulaga

Technical Team

Pedro Agostinho
Rafal Kulaga
Antonin Kveton
Ewald Sperrer
Martin Zemko

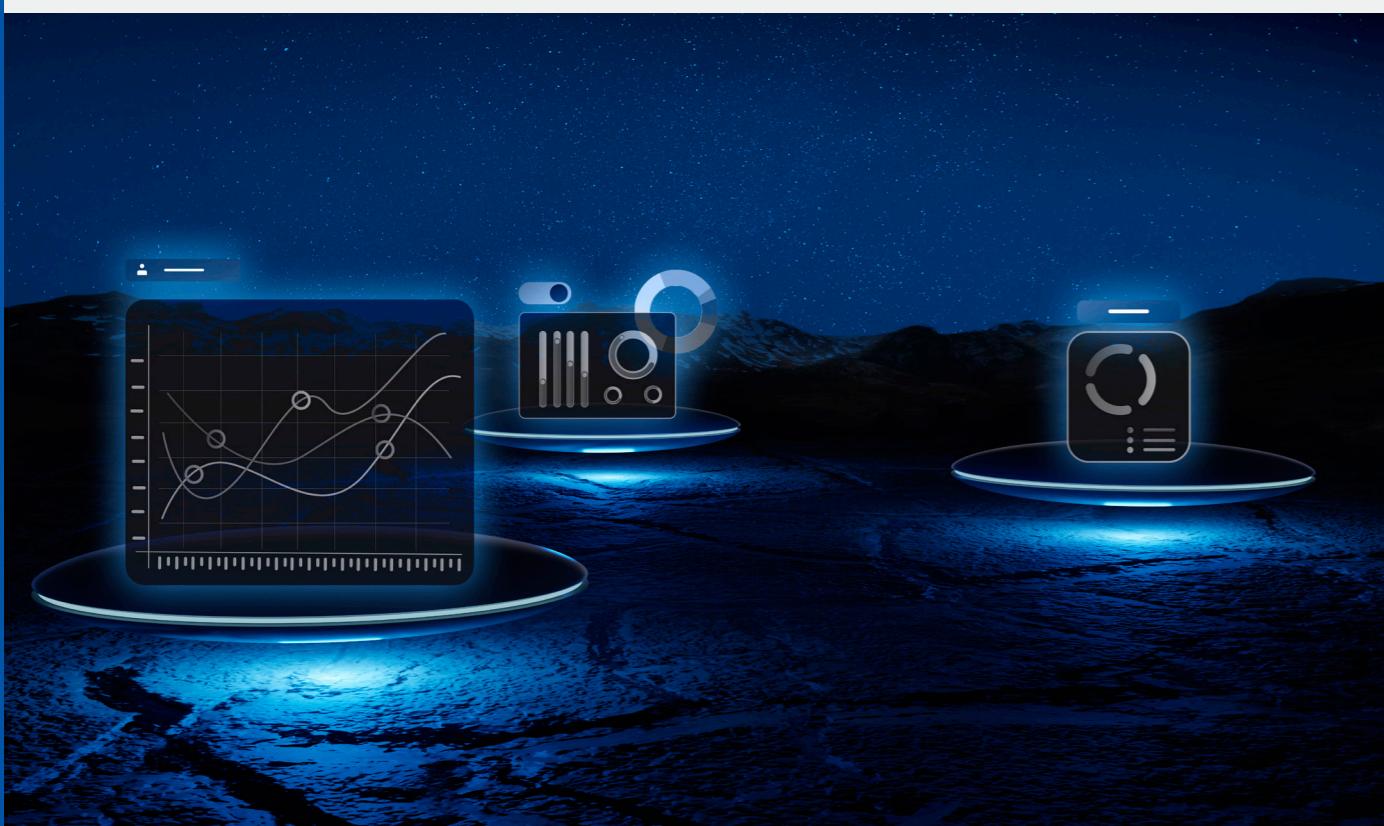
Collaboration Liaisons

Pedro Agostinho
Ewald Sperrer
Christopher Stoegerer

SIEMENS

Data Analytics for Industrial Control Systems

Aims to enhance the efficiency, reliability, and intelligence of the industrial control systems used within CERN's accelerator complex. A primary objective is to develop a scalable device monitoring solution by creating a web-based application prototype that leverages edge computing technologies and real-time data analytics to monitor control devices. Additionally, the project seeks to assess and adapt Siemens' industrial solutions to meet the specific operational needs of CERN. Another significant goal is to explore and evaluate the potential of open-source Large Language Models (LLMs) for applications in control systems, aiming to enhance engineering task automation, troubleshooting, and decision-making processes.



Overview

The upcoming High-Luminosity LHC (HL-LHC) upgrade is poised to increase the particle collision data sample by a factor of ten compared to the current LHC program. This significant enhancement will inevitably lead to increased complexity in the associated control systems. Consequently, improving the functionality and scalability of these systems is essential to meet future operational demands. This task involves several interconnected initiatives. One subproject focuses on developing a robust device monitoring platform to manage the hardware components of industrial control systems. This platform will utilize edge computing technologies and real-time data analytics to ensure reliability. Another key subproject involves evaluating open-source Large Language Models (LLMs) for potential applications in control systems. The goal is to explore how LLMs can enhance task automation, streamline diagnostics, and assist decision-making processes.

Highlights in 2024

In 2024, several key milestones were accomplished in the project. A functional prototype of the web application for Device Monitoring was developed at CERN, featuring both frontend and backend components. The front end enables end-users to configure and organize control devices into a hierarchical, tree-like structure while allowing them to design logical rules for individual categories or tree nodes. These rules are executed in real-time, playing a vital role in hierarchically displaying the status of the entire control system. This real-time visualization simplifies navigation and enables quick identification of system errors. The backend component was designed to handle rule execution efficiently. The application also supports third-party integrations, including compatibility with Siemens' device monitoring solutions, such as Machine Insight. Additionally, as part of another subproject, we explored continual pre-training of Large Language Models (LLMs) to acquire domain-specific knowledge from industrial control documentation. A prototype benchmarking suite was also developed to evaluate various LLMs on standardized datasets.

Next Steps

The project is set to continue in the coming years, with certain subprojects potentially shifting to new focus areas based on priorities determined in collaboration with Siemens. These focus areas may include the exploration of virtual Programmable Logic Controllers (PLCs) deployable on industrial edge devices, the study of predictive maintenance solutions for control devices using historical data, and the investigation of small language models for targeted industrial applications. A workshop between CERN and Siemens is planned to refine specific use cases and define objectives for the years ahead.

Project Coordinator

Fernando Varela Rodriguez

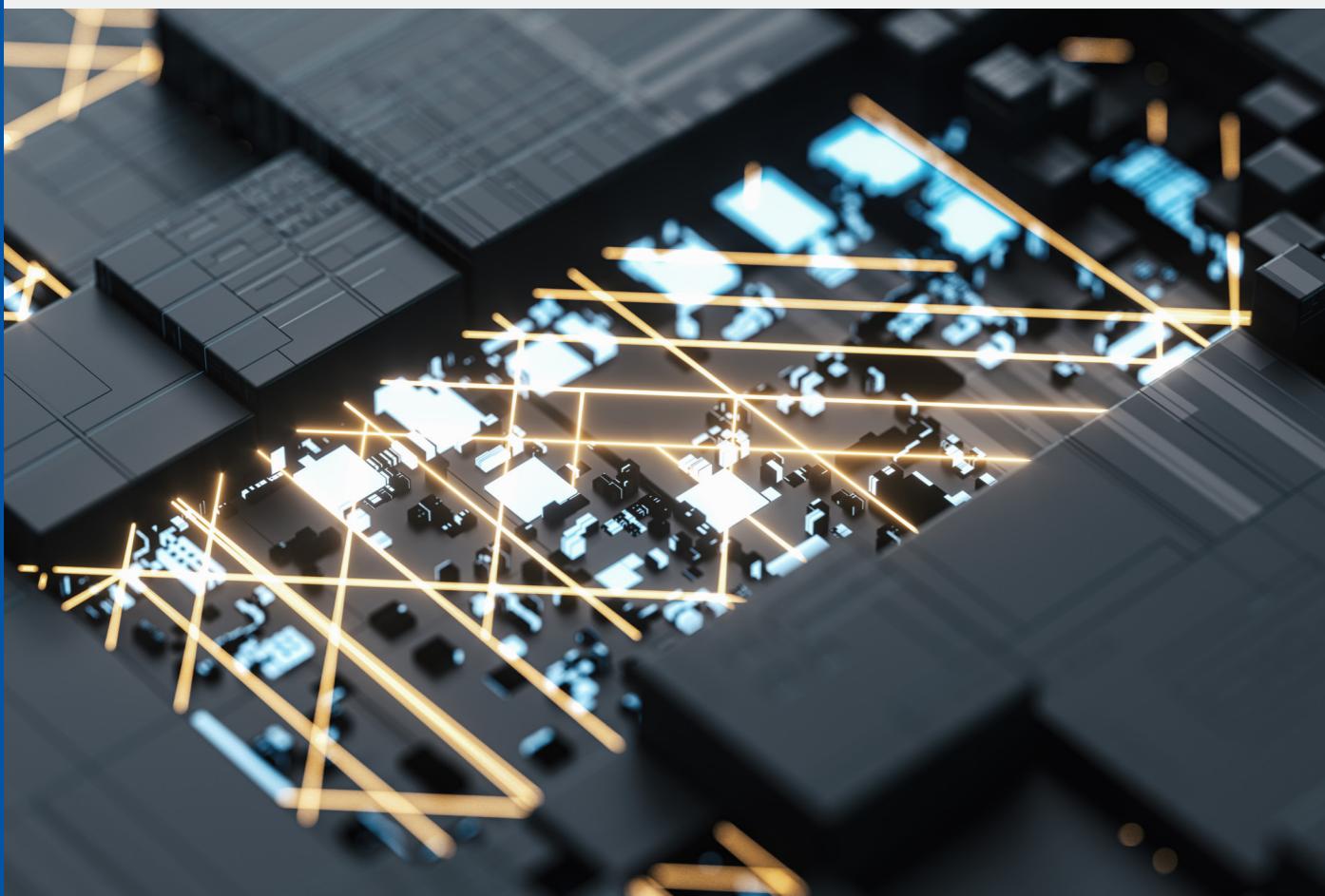
Technical Team

Aabit Patil
Fernando Varela Rodriguez
Jeronimo Ortola Vidal

Collaboration Liaisons

Silvio Becher
Christian Kern
Stefan Langer

SIEMENS



Heterogeneous Architectures Testbed

Aims to provide a diverse hardware portfolio for comprehensive technology testing. Focused on assessing the efficacy of various architectures, this initiative aims to provide valuable insights into the practical utility of emerging technologies. By subjecting a spectrum of hardware configurations to real-world applications, the project seeks to establish benchmarks that guide the adoption of the most effective and efficient technologies.

Overview

The project is pivotal for CERN and broader scientific communities. By creating a diverse portfolio of applications tailored for various hardware architectures, the project aims to enhance technology evaluation. This work is crucial for optimizing computational efficiency, fostering innovation, and ensuring that CERN and the wider scientific community stay at the forefront of technological advancements, ultimately advancing our capabilities in high-performance computing and scientific research.

Highlights in 2024

The project has achieved significant milestones in evaluating cutting-edge technologies across diverse CPU platforms, including the latest x86 (Intel® Xeon®, AMD EPYC™) and Arm (Ampere® Altra®, NVIDIA Grace™) CPUs as well as GPUs such as Intel®Flex, Intel GPU Max and NVIDIA H100. The project extensively ran benchmarks with a focus on energy consumption measurements, providing a comprehensive assessment of efficiency.

Notable example applications and benchmarks such as HEPScore23 and MadGraph were thoroughly tested. Additionally, the testbed allowed extensive software testing for the CMS, ATLAS, and LHCb experiments. The support continuously given to the user community emerged as a crucial aspect, possibly the most important, fostering collaboration, offering valuable feedback, and ensuring seamless integration of new hardware.

This progress played a pivotal role in fostering collaboration with industry partners, offering valuable feedback on technology performance and shaping future developments. The project's outcomes are integral not only for advancing CERN experiments but also for guiding industry stakeholders in optimizing their technologies for real-world applications. The collaborative and open approach ensures seamless integration of new hardware, enhancing computational capabilities and fostering groundbreaking advancements in scientific research.

Next Steps

The project's next steps involve ongoing evaluation of emerging technologies, crucial for readiness in the High-Luminosity LHC era. Continuous assessment of advancements in x86, Arm and RISC-V CPUs, as well as upcoming GPU platforms, remains a priority. Additionally, the investigation includes the potential integration of remote data centres, high-performance computing (HPC) centres, and public commercial cloud resources. This forward-looking strategy ensures that the "Heterogeneous Architectures Testbed" stays at the forefront of technological innovation to meet the computational demands of the evolving scientific landscape.

Project Coordinator

Luca Atzori

Technical Team

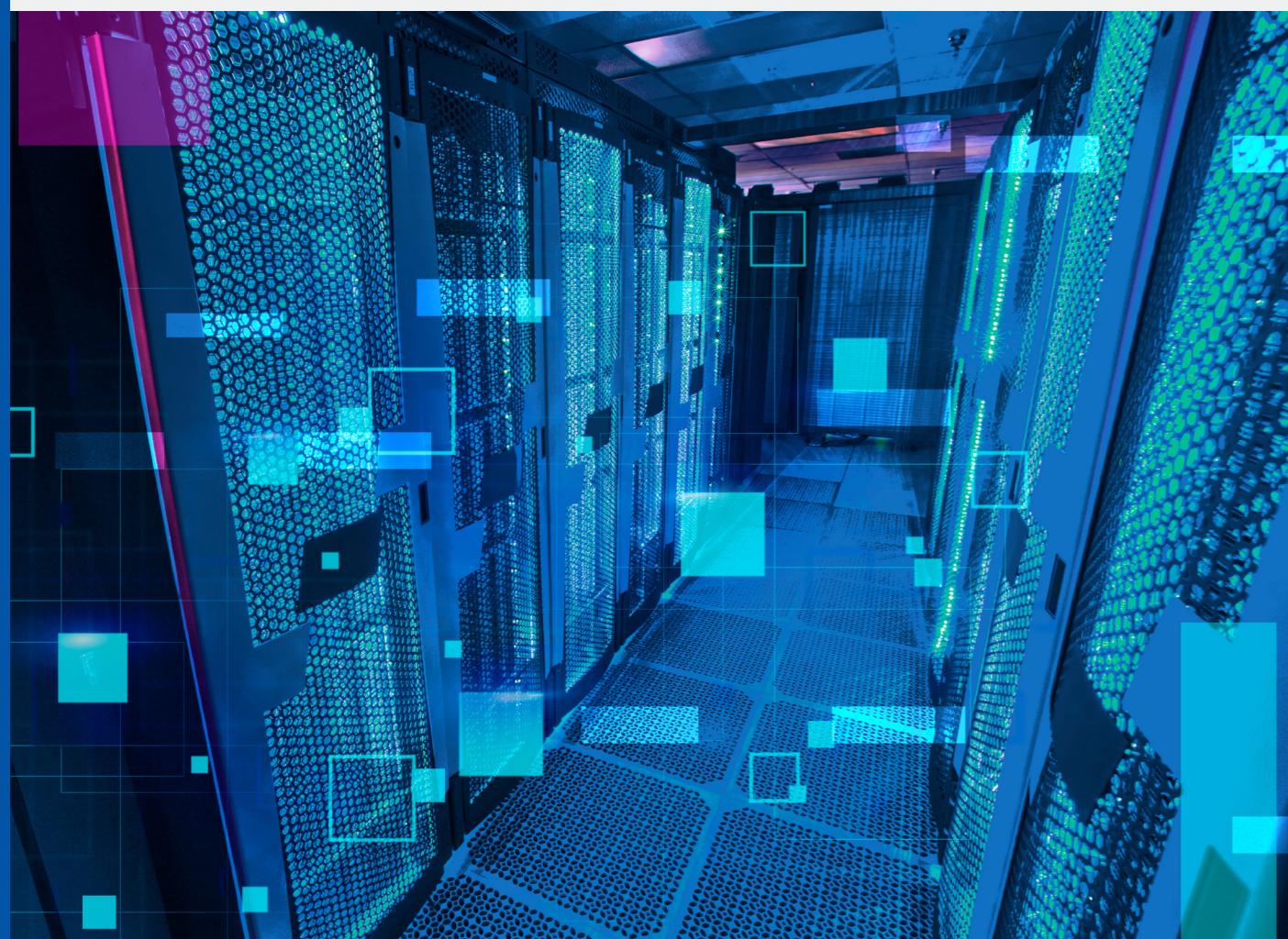
Luca Atzori
Albane Carcenac
Maria Girone
Joaquim Santos
Jessy Sobreiro
David Southwick
Eric Wulff

Collaboration Liaisons

Claudio Bellini
Marco Cicala
Ian Fisk
Cosimo Gianfreda
Daniele Gregori
Bruno Riva
Walter Riviera

E4
COMPUTER
ENGINEERING

SIM NS
FOUNDATION



SPECTRUM

The EU-funded SPECTRUM project unites top European science organizations and e-Infrastructure providers to develop a Strategic Research, Innovation, and Deployment Agenda (SRIDA) and a Technical Blueprint for a European compute and data continuum. This initiative aims to establish an Exabyte-scale research data federation and compute continuum, enhancing data-intensive scientific collaborations across Europe.

Overview

Frontier research in High-Energy Physics and Radio Astronomy is entering the Exascale era, with new instruments requiring unprecedented data processing capabilities. To meet these demands, pan-European data and compute infrastructures must be developed, incorporating novel architectures, federation models, and IT frameworks. The integration of Exascale HPC and Quantum computing systems offers new opportunities to accelerate discoveries and complement existing research facilities. Key challenges include scalability, performance, energy efficiency, portability, interoperability, and cybersecurity, which must be addressed to ensure successful integration of these heterogeneous systems.

Highlights in 2024

In 2024, the SPECTRUM project made significant progress in establishing its organizational structure and initiating key activities. Following a strategic kick-off meeting, the project launched the SPECTRUM Community of Practice (SPECTRUMCoP), complete with a published charter and active Working Groups. Work Packages 3 and 5 led efforts in developing use-case analysis frameworks and conducting a comprehensive survey on current and future needs. The project demonstrated strong inter-WP collaboration, particularly in structuring the SPECTRUMCoP and refining use-case analyses. With the Activity Management Board overseeing operations and CERN contributing to multiple Work Packages, SPECTRUM has laid a solid foundation for its ambitious goal of creating an Exabyte-scale research data federation and compute continuum across Europe.

Project Coordinator

EGI Foundation

Technical Team

Maria Girone, David Southwick, Eric Wulff

Collaboration Liaisons

Sergio Andreozzi, Xavier Salazar, Chiara Ferrari, Tommaso Boccali, Hans-Christian Hoppe, Corentin Lefevre, Raymond Oonk, Fabio Affinito, John Swinbank

Next Steps

SPECTRUM will move forward, in the coming months, with several key activities:

Analysis of survey results: The project team will analyze the responses from the survey to provide valuable insights to guide future project directions.

Refinement of use-case analyses: The project will focus on developing and refining additional representative use cases. This will help in understanding the specific requirements and challenges across different research domains.

Continuation of Working Group activities: The established Working Groups (WGs) within the SPECTRUM Community of Practice (SPECTRUMCoP) will continue their bi-monthly meetings.

Preparation for upcoming Work Packages: Work Packages 4, 6, and 7 are scheduled to start in later months of the project. The team will begin preparations for these upcoming activities.

Development of the Strategic Research, Innovation, and Deployment Agenda (SRIDA): Based on the insights gathered from surveys, use-cases, and Working Group discussions, the project will work towards formulating the SRIDA.

Progress on the Technical Blueprint: Alongside the SRIDA, the project will continue developing the Technical Blueprint for the European compute and data continuum.

These next steps will contribute to SPECTRUM's overarching goal of creating an Exabyte-scale research data federation and compute continuum to enhance data-intensive scientific collaborations across Europe.



SPECTRUM is funded by the European Union Grant Agreement Number 101131550

Center of Excellence on AI and Simulation-Based Engineering at Exascale (CoE RAISE)

CoERAISE is an EU funded project and CERN leads Work Package 4 (WP4) which aims at the development and expansion of AI methods along representative use-cases from research and industry, which have a strong focus on data-driven technologies, i.e., analyzing data-rich descriptions of physical phenomena. The outcomes are applicable to intelligent workflows including innovative AI methods and techniques, optimized on HPC-to-Exascale systems. The tasks contain the capabilities to evaluate prototype algorithms based on experimental and/or simulation data, code performance on Exascale HPC systems, and quality of data models.



Overview

CERN leads WP4 which contains four tasks:

Event reconstruction and classification at the CERN HL-LHC, led by CERN (T4.1)

Seismic imaging with remote sensing for energy applications, led by the Cyprus Institute (T4.2)

Defect-free additive manufacturing, led by Flanders Make (T4.3)

Sound Engineering, led by University of Iceland (T4.4)

Task 4.1 develops a GPU native and AI-based algorithm for particle-flow reconstruction that can easily be accelerated by modern heterogeneous hardware. This algorithm, called Machine-Learned Particle-Flow (MLPF), is developed in collaboration with CMS and acts as a representative AI use case from HEP. Some of the most important contributions from T4.1 include the implementation and execution of distributed training and large-scale hyperparameter optimization using HPC, significantly improving physics performance. Another area of work has been to optimize developed algorithms on various heterogeneous architectures.

Project Coordinator

Maria Girone
Andreas Lintemann

Technical Team

Maria Girone, David Southwick, Eric Wulff

Collaboration Liaisons

Marcel Aach, Naveed Akram, Gabriele Cavallaro, Kurt de Grave, Arnis Lektuaers, Andreas Lintemann, Morris Riedel, Nikos Savva, Eric Michael Sumner, Liang Tian, Eric Verschuur

Highlights in 2024

In T4.1, the MLPF (Machine-Learned Particle Flow) studies on the open electron-positron collision dataset has been completed and focus has shifted back to simulated CMS-based datasets with proton-proton collisions. A strategic decision was made to migrate the optimization code of MLPF from TensorFlow to PyTorch. The reason for this is the superior support for cutting edge tools offered by PyTorch as well as its suitability for easy and fast development of new algorithms. The work started in 2023 and has now been completed. To further test the exascale potential of AI workloads like MLPF, a study on the neural scaling laws of MLPF was carried out using the new PyTorch version of the model and training code. This study was presented at the PASC24 conference in Zurich, Switzerland.

The CoE RAISE project came to an end in July of 2024 and the final EC review was carried out in September. The project received nothing but good feedback from the reviewers, who congratulated all project partners and praised the project outcomes.

Next Steps

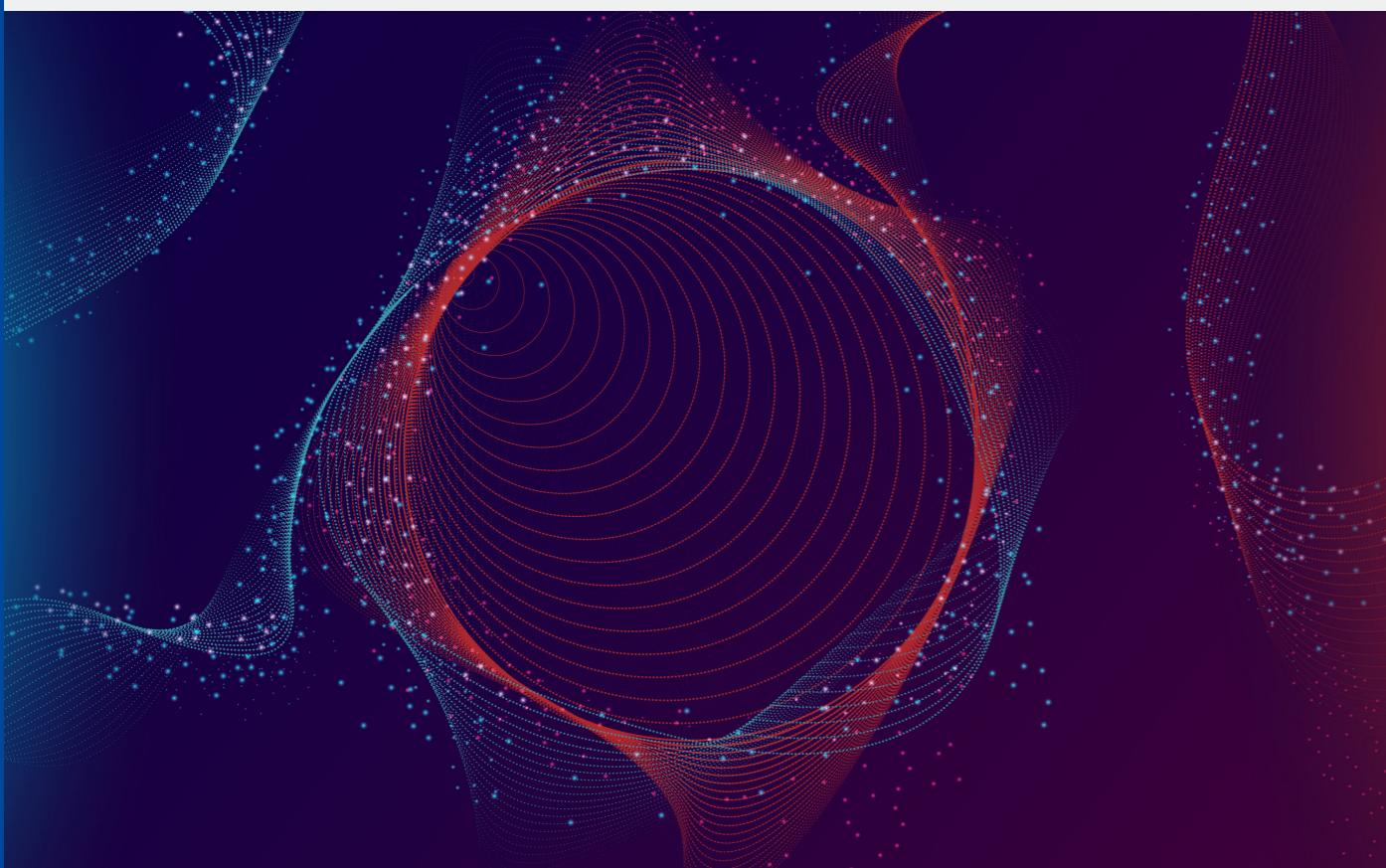
CoE RAISE concluded its activities at the end of July 2024. Following this, the MLPF group will continue its work on AI-based particle flow reconstruction. However, this ongoing effort will experience a reduction in support from CERN openlab due to the cessation of funding associated with CoE RAISE.



The CoE RAISE project have received funding from the European Union's Horizon 2020 – Research and Innovation Framework Programme H2020-INFRAEDI-2019-1 under grant agreement no. 951733

Quantum Databases for Dynamic Data Storage

Exploring the usage of a quantum database with potential quantum indexing and classical or quantum data elements that may be initially of unknown length. By classical, we mean specifically cloneable states in terms of the no-cloning theorem. Thus, we focus on creating an algorithmic procedure for efficiently manipulating such states in quantum software and exploring its use cases and applications.



Key elements in quantum mechanics include the presence of superposition states and entanglement. In particular, the superposition principle allows access to data in a coherent way. Thus, within the concept of a quantum database, an algorithmic framework for efficiently manipulating structured data in such a superposition state is provided. In our project, we are, for example, concerned with data obtained through experiments over an unknown temporal interval. Thus, such experimental data can be inherently of non-predefined length, e.g., the runtime of the experiment is not given initially. Hence, we are mainly concerned about its resource-efficient data manipulation methods and the specific data operations similar to classical databases that are applied within a dynamical scenario in quantum software.

Highlights in 2024

In 2024, we developed a set of efficient algorithms that implement the protocols mimicking classical database operations in the quantum scenario with respect to quantum indexing and explored their feasibility. This is relevant for future use cases and may serve as a guideline to define which classical database operations are, in general, useful as soon as quantum indexing or quantum data is present. The algorithms were formally defined and implemented as a proof of concept in Python and C++ using the Intel Quantum simulator.

Next Steps

The project's next steps will be to focus on various applications and precise use cases.

Project Coordinator

Michele Grossi, Sofia Vallecorsa

Technical Team

Michele Grossi, Gian Giacomo Guerreschi, Carla Sophie Rieger, Sofia Vallecorsa, Martin Werner

Collaboration Liaisons

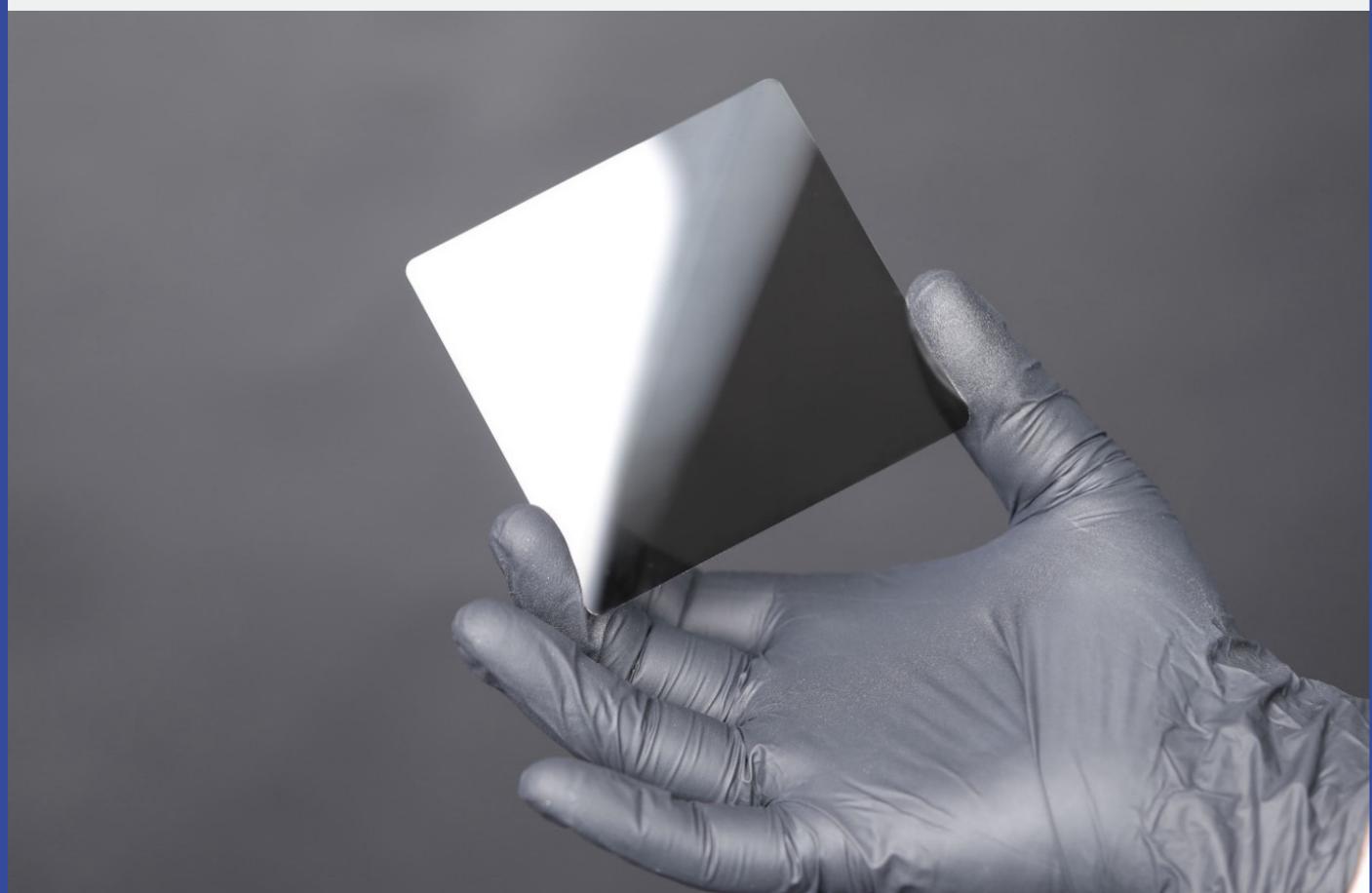
Gian Giacomo Guerreschi, Martin Werner

Emerging Technologies



Evaluation of Cerabyte: Archival Data Storage Technology using Ceramic Nanolayers

This project will allow CERN to gain understanding of the possibilities and potential problems of the new archive storage medium using ceramic nanolayers. It represents an opportunity to provide feedback and to guide the development of this technology in a way that is beneficial to CERN.



Overview

Currently tape is CERN's archival storage medium of choice, due to its characteristics of reliability, durability and price. While the roadmap for tape is encouraging, tape as a technology is exposed to certain market risks, and CERN should always have an eye on the development of possible alternatives and likely evolution of archival storage.

For many years, IT-SD group has actively engaged with companies doing R&D in innovative storage technologies, including DNA, archive glass and ceramic. At this point, ceramic seems to be the most likely contender to achieve sufficient I/O rates to be useful, within a reasonable time frame (~10 years). The success of this technology would represent a disruptive change to the landscape.

Project Coordinator

Vladimir Bahyl

Technical Team

Vladimir Bahyl

Collaboration Liaisons

Ed Childers
Vladimir Bahyl
Sebastian Kirsch
Martin Kunze

Highlights in 2024

2024 was the first year of the collaboration. CERN benefited from the in-person participation of Cerabyte at the openlab workshop (in March) to explain the current and future data storage needs. Reciprocally, Cerabyte outlined their product development roadmap. The collaboration was officially announced in August via a BusinessWire press release.

The collaboration has been influential to the Cerabyte as during the course of discussions CERN has made some product recommendations for consideration. Cerabyte has adopted CERN's recommendations.

Cerabyte is developing a digital data format on glass with a robust ECC design as an alternative to QR codes currently used in the prototype system.

Cerabyte is pursuing a cartridge design for the glass carriers which will leverage the mature infrastructure of the LTO tape cartridge ecosystem.

In March, at CERN's invitation, Cerabyte presented about "Long-term high-volume data storage in ceramic nano layers" at ACAT 2024

In June CERN participated in an investor reference call and in July CERN provided 1 GB of sample data which will be written to a data carrier.

Next Steps

The collaboration will focus on 3 areas in the coming year:

Validation of the data format. CERN's 1GB of real world data will be written with the new ECC format. CERN's assistance in validation of the carriers in stress/environmental conditions will be very valuable.

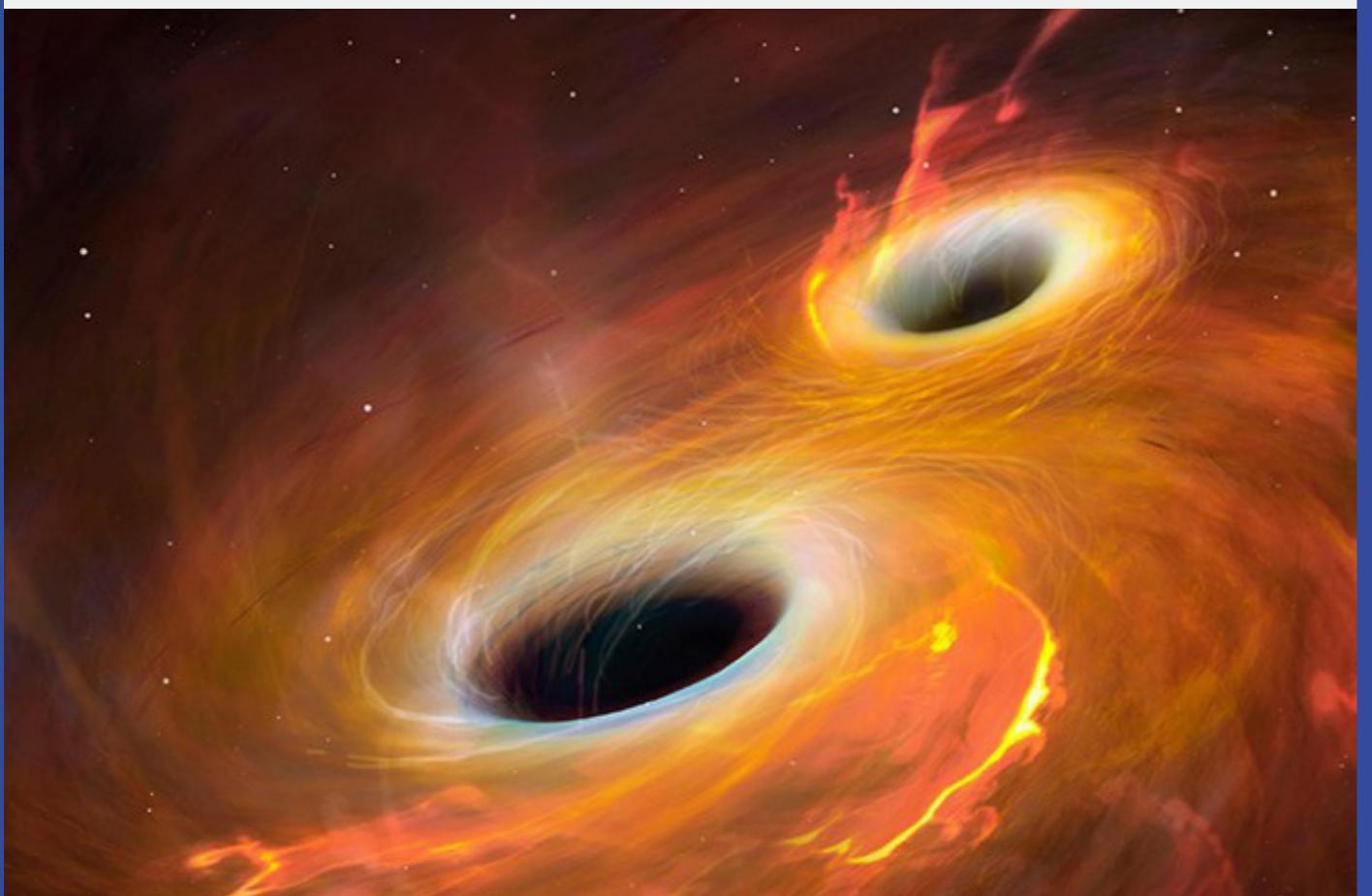
Cerabyte would like to work with CERN toward establishing an open format standard for the Cerabyte data carrier.

Cerabyte would like to use CERN's experience in the development of a self-describing format for data written to glass carriers.



interTwin: Co-designing and Prototyping an Interdisciplinary Digital Twin Engine

interTwin seeks to co-design and implement a prototype of an open-source Digital Twin Engine (DTE), built upon open standards, facilitating seamless integration with application-specific Digital Twins (DTs). This innovative platform, rooted in a co-designed interoperability framework and the conceptual model of a DT for research, known as the DTE blueprint architecture, aims to simplify and accelerate the development of complex application-specific DTs.



Overview

InterTwin develops and implements an open-source DTE that offers generic and customized software components for modeling and simulation, promoting interdisciplinary collaboration. The DTE blueprint architecture, guided by open standards, aims to create a common approach applicable across scientific disciplines. Use cases span high-energy physics, radio astronomy, climate research, and environmental monitoring. The project leverages expertise from European research infrastructures, fostering the validation of technology across facilities and enhancing accessibility. InterTwin aligns with initiatives like Destination Earth, EOSC, EuroGEO, and EU data spaces for continuous development and collaboration.

Highlights in 2024

The itwinai toolkit saw a stable release this year, supporting AI workflows on cloud and HPC with integration of distributed ML training frameworks, ML tracking systems, and modular workflows. Extensive documentation and tutorials were developed, covering Distributed Machine Learning, ML workflows, and scaling benchmarks. We also completed infrastructure integration with interLink and OSCAR for remote ML inference on HPC and interactive model development on JupyterLab for Vega. Significant progress was made in integrating CERN and other use cases into the digital twin engine, covering both physics and climate research domains.

In 2024, significant progress was made on CERN's use case tasks within the InterTwin project. Efforts focused on refining and updating the generative model (3DGAN) analysis and validation frameworks. Key activities included updating the codebase to ensure compatibility with ML framework requirements, particularly regarding tensor operations and data types. Model validation scripts were developed and refined, with integrated training experiments conducted utilizing the itwinai framework on JSC resources. Contributions included status inputs for project deliverables, presentations at the 4th Technical Meeting, summer student lectures, and the 2nd EC project review.

Next Steps

Next, we will consolidate itwinai for containerized workflows on cloud and HPC to complete the integration with the workflow manager of the Digital Twin Engine, and fully implement the hyper-parameter optimization module. Additional integrations are planned for radio astronomy, lattice quantum chromodynamics, EURAC environmental data, and climate modeling.

The next steps for CERN's use case into the InterTwin project involve several targeted actions to build on the progress achieved in 2024. A primary focus will be finalizing the performance validation framework and possible integration with DTE core modules that further enable our DT's capabilities. Additionally, the team plans to integrate a model from the CaloChallenge under the detector simulation use case into the itwinai framework, enabling more robust and scalable simulations. This integration will be accompanied by continued refinement of model validation scripts to validate and benchmark model performance. Further dissemination activities are also anticipated, including contributions to technical meetings and collaborative publications.

Project Coordinator

EGI Foundation

Technical Team

Matteo Bunino, Anna Elisa Lappe, Xavier Espinal, Enrique Garcia, Maria Girone, Jarl Saether, Kalliopi Tsolaki, Sofia Vallecorsa

Collaboration Liaisons

Donatello Elia, Sandro Fiore, Vera Maiboroda, David Rousseau



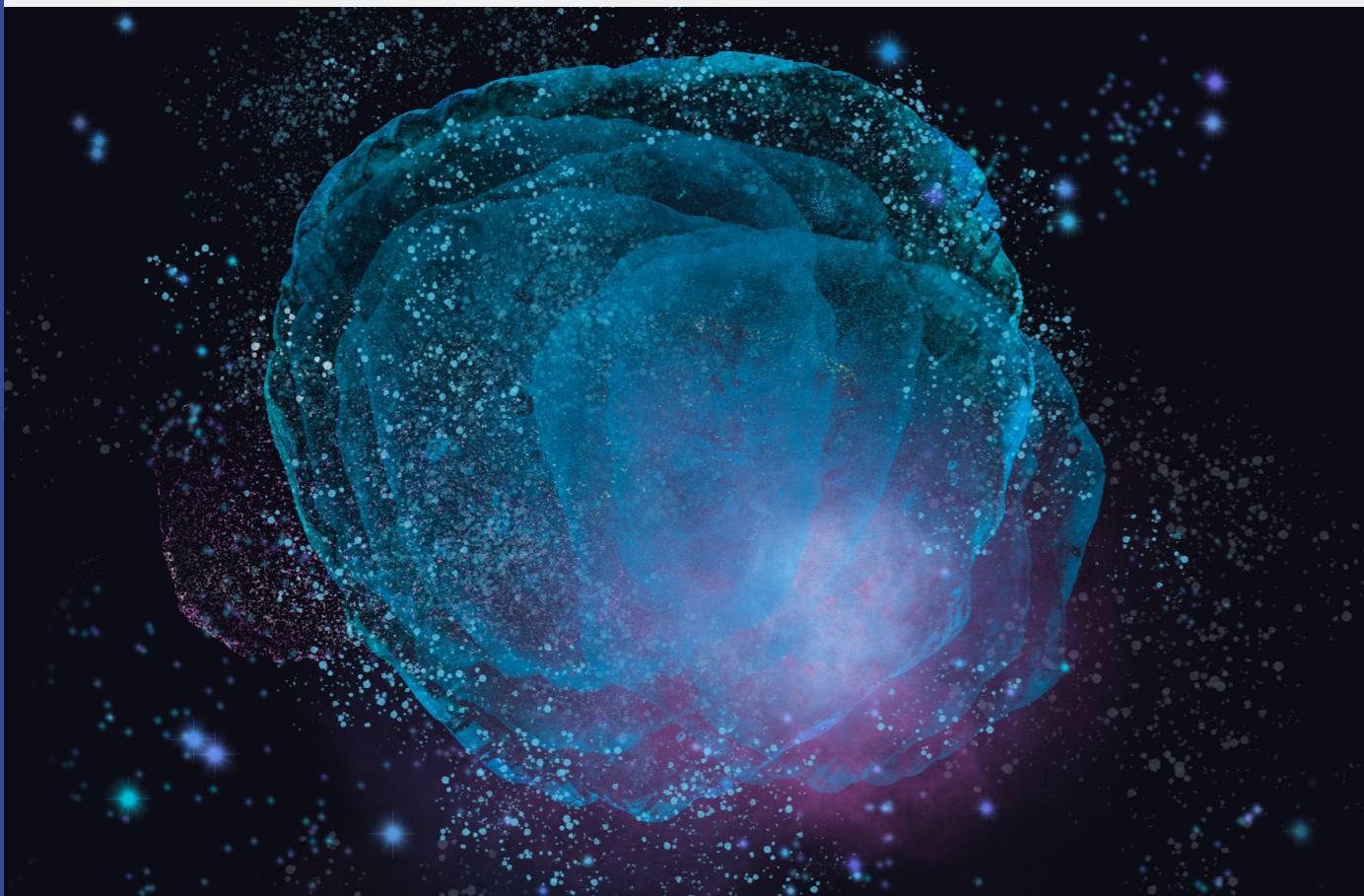
interTwin

interTwin is funded by the European Union
Grant Agreement Number 101058386



BioDynaMo: Biology Dynamics Modeller

BioDynaMo aims to develop a platform that empowers scientists to effortlessly generate, execute, and visualize agent-based simulations. Utilizing cutting-edge computing technologies, the BioDynaMo platform will facilitate simulations of unprecedented scale and complexity. This capability opens avenues for addressing intricate scientific research inquiries with greater ease.



Overview

In the life sciences community, computer simulation is gaining prominence for modelling intricate biological systems. While numerous specialized tools exist, creating a high-performance, versatile platform represents a significant advancement. CERN leverages its extensive expertise in large-scale computing, supported by funding through the CERN and Society Foundation, to address three highly relevant societal needs: the fight against cancer, inequalities and dengue. Additionally, BioDynaMo is part of the CERN IT department's efforts to build a Digital Twin Engine for Science, with its integration in projects such as interTwin, making it accessible to a wider range of researchers.

Highlights in 2024

In October 2024, our work on 'Calibration of stochastic, agent-based neuron growth models with approximate Bayesian computation' was published on the Springer Journal of Mathematical Biology. This work established a robust framework for calibrating agent-based neuronal growth models and opens the door for future investigations using Bayesian techniques for model building, verification, and adequacy assessment.

There were also significant advancements in developing the distributed, MPI-parallel platform, with a publication expected during 2025.

By the end of 2024, a project led by the GSI Biophysics Department, using the BioDynaMo simulation platform, was awarded a place among the Top 10 scientific breakthroughs of 2024, by Physics World. This recognition highlights the project's ground-breaking contribution to physics and computational biology, specifically in modelling lung cells to personalize radiotherapy—a leap forward in precision oncology.

Next Steps

Our plan is to make significant strides in enhancing the overall usability of our platform, with a dedicated focus on refining user experience and robustness. We will actively extend our documentation to ensure clarity and accessibility. Furthermore, we will continue our commitment to advancing various use cases in biology and beyond, continuously exploring innovative solutions and contributing to the broader scientific community.

Project Coordinator

Maria Girone

Technical Team

Lukas Breitwieser
Tobias Duswald,
Maria Girone
Stavros Portokalidis
Eric Wulff

Collaboration Liaisons

Roman Bauer
Vasileios Vavourakis



EMP2: Environmental Modelling and Prediction Platform

EMP2 aims to develop a proof-of-concept for a machine learning based digital twin of the atmosphere for environmental applications. To accomplish this, the project is subdivided into two main parts. The first segment will focus on the development of a machine learning based modelling core prototype, called AtmoRep, built on the concept of large scale representation learning applied to Earth System Science. In the second phase, the modelling core will be integrated into the digital twin architecture currently under development within the CERN IT department by the InterTwin project.



Overview

The atmosphere and its dynamics have a significant impact on human well-being, from agricultural decision making, to policy making and the renewable energy sector. An accurate and equitable modeling of atmospheric dynamics is consequently of critical importance to allow for evidence-based decision making that improves human well being and minimizes adverse impacts for current and future generations. Very recently, AI-based models have shown tremendous potential in reducing the computational costs for numerical weather prediction. However, they lack the versatility of conventional models. The EMP2/AtmoRep project aims at developing an AI-based model of atmospheric dynamics for multi-purpose applications. The model has been implemented leveraging the concept of large-scale representation learning, so to encapsulate the information from the large amounts of available data. The implementation on the digital twin platform will make such information more accessible to the general public, allowing the users to easily develop their own applications in weather and climate.

Highlights in 2024

In 2024 we published the model on GitHub, making it fully open source. During summer we focused in getting the model smaller and more efficient. We moved from a 3.5B parameter model, to a 600M parameter core model which is 6 times faster and 30% more accurate than the previous published architecture. In addition, we have started implementing the roll-out mechanism to enable medium range global weather forecasting up to 10-15 days. Through the HClimRep collaboration, kicked-off in September 2024, the model is getting upgraded towards decadal time scales predictions. The latest changes include the integration of ocean and stratosphere level information into the current architecture. In the last part of the year, in collaboration with interTwin, we have settled a roadmap for the integration of the model within the interTwin digital twin platform for AI models, which will be completed in January 2025.

Next Steps

The project will end in February 2025. The model is already serving as the basis for the HClimRep project in the context of the Helmholtz AI call on foundation models. Starting from February 2025, the developed core model will serve as basis for the development of the WeatherGenerator, an EU funded project accounting for 16 partners and led by ECMWF. The goal of the WeatherGenerator will be to build a foundation model for a large number of tasks, from renewable energy and flood prediction to food security, health and the biosphere. In 2025 the model will be used in the context of the WFP-LIST-CERN strategic partnership on AI, as atmospheric model for seasonal predictions of crop yields in critical areas of the globe.

Project Coordinator

Alberto Di Meglio
Ilaria Luise

Technical Team

Alberto Di Meglio
Ilaria Luise

Collaboration Liaisons

Michael Langguth
Christian Lessig
Martin Schultz



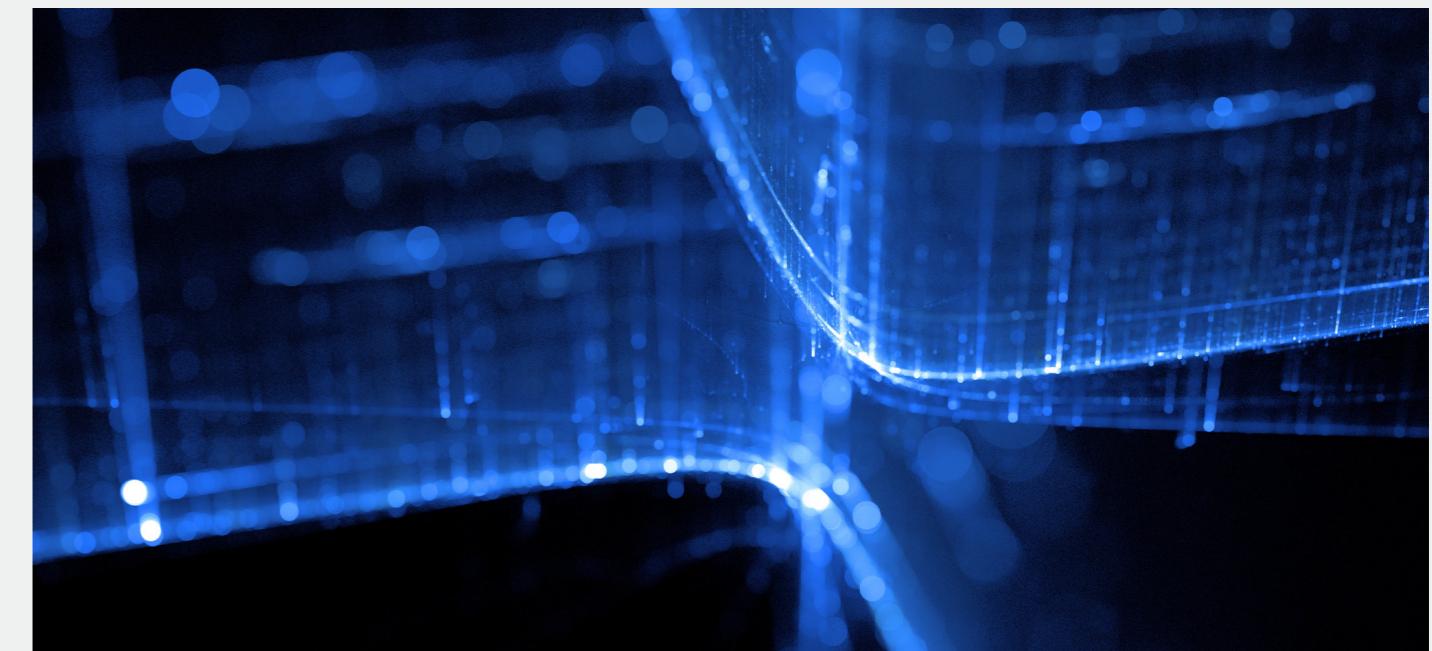


Starting in 2025

Next-Generation Exascale Flash Storage

Aims to leverage the expertise of both partners in storage system design. CERN plans to evaluate and integrate the latest flash storage technologies from Pure Storage into its large-scale distributed storage system. This collaboration aims to optimise performance, ensure seamless scalability, and address future scientific challenges. Pure Storage R&D resources will partner with CERN to co-design highly efficient solutions, maximise the value of deployed technologies, and refine product development to better meet the evolving needs of CERN, as well as other communities and industries.

In partnership with:

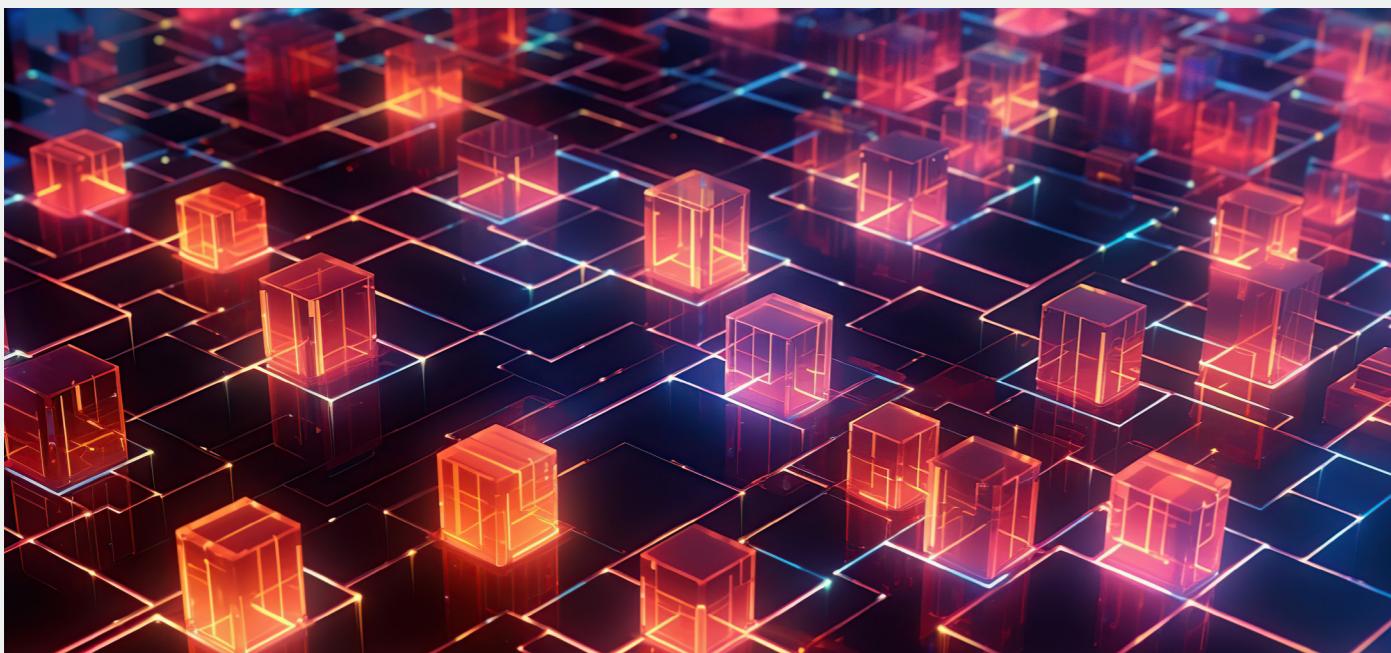


Oracle Kubernetes Operator

Aims to use the Oracle Kubernetes Operator and other Cloud Native Computing Foundations tools to modernise the Oracle Database and Oracle REST Data Services (ORDS) services. The Oracle Operator will help to automate database provisioning, configuration, and management, simplifying operations with simple Kubernetes resources. It also aims to facilitate the adoption of GitOps techniques for the ORDS infrastructure (already hosted on Kubernetes). Additionally, the Oracle Operator will complement tools like Crossplane, supporting the implementation of Infrastructure as Code for infrastructure management.

In partnership with:

ORACLE



Cost Optimization and Sustainability for Public Cloud Provider

Aims to identify and implement metrics to understand the carbon footprint of the current CERN services hosted on Kubernetes. These metrics will inform the deployment of workloads in function of the Scope 3 emissions across on-premises and public clouds as a path to achieve low/zero-carbon workload deployment. Ultimately, these innovative results will have direct impact in the IT Department cloud operating model, where carbon awareness will be a costing factor being considered in the FinOps activity managing public cloud costs and optimizing cloud spending within CERN.

In partnership with:

ORACLE

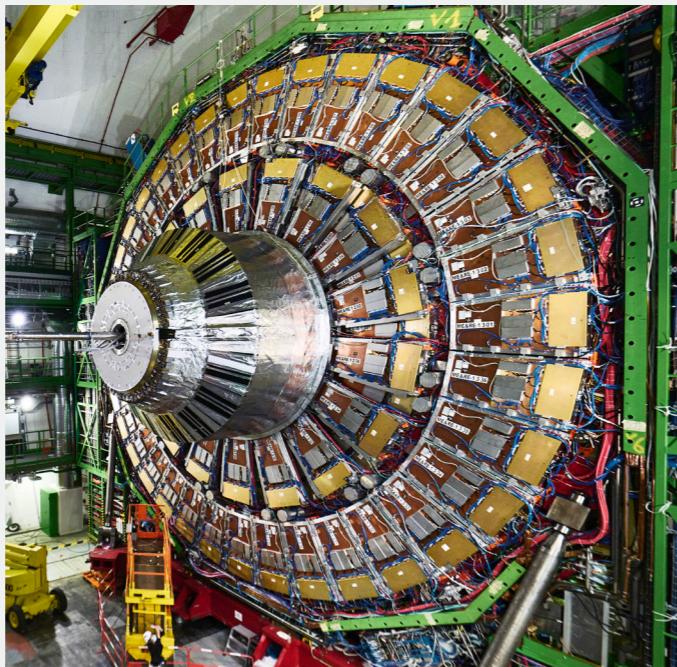
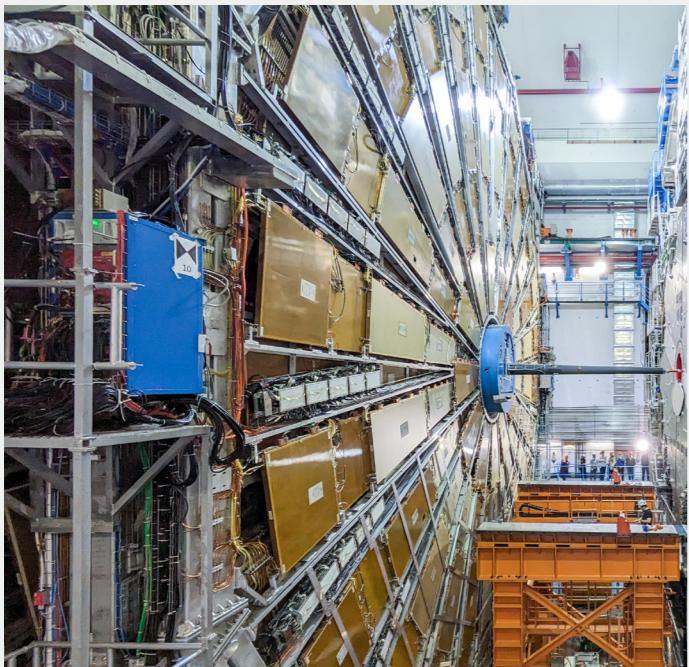


Anomaly Detection for Ultra Low Latency Event Selection at the LHC

Aims to implement anomaly detection on FPGA devices in order to look for physics that may have been missed with standard trigger techniques. Tools such as high level synthesis languages, and ML accelerator engines will be deployed to facilitate the implementation when advantageous. The CMS and ATLAS experiments will work together to share experience and pitfalls in order to make use of this solution already before the end of Run-3.

In partnership with:

ORACLE

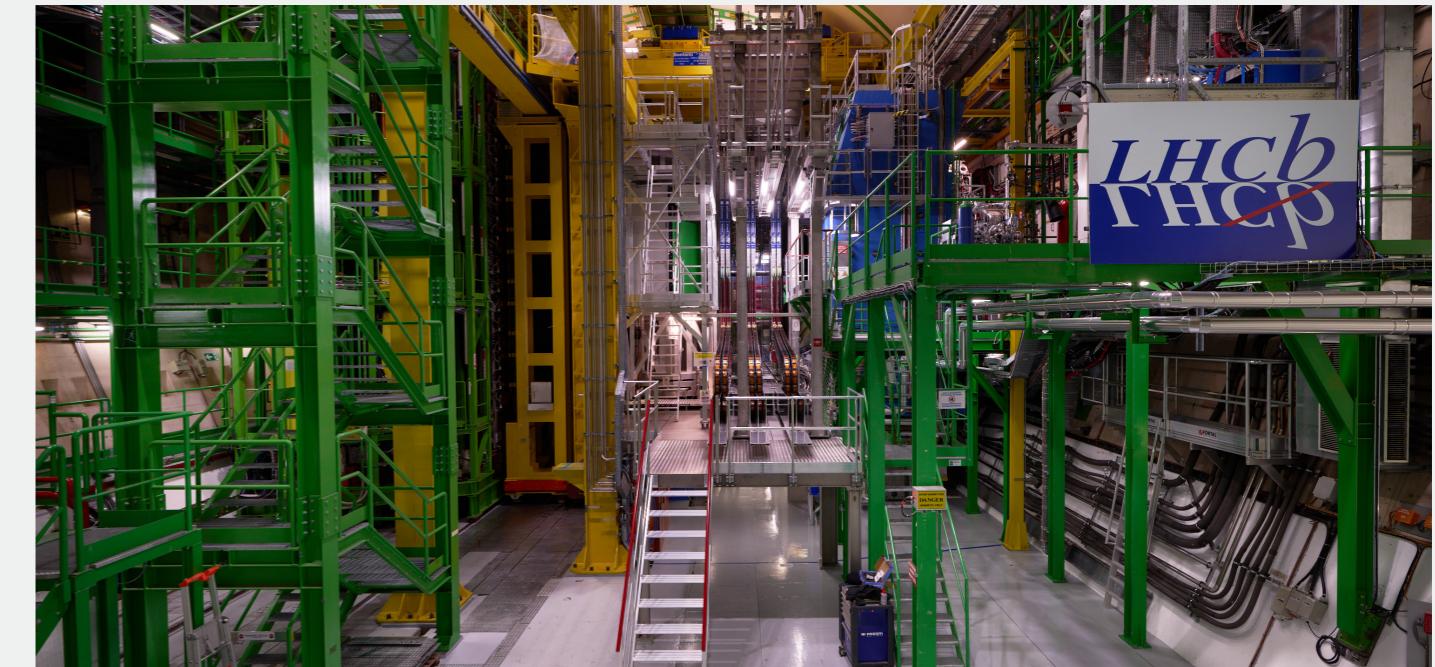


Digital Twin: Data Science Engine

The project aims to address critical gaps in data analysis and predictive modeling by employing advanced mathematical models and machine learning techniques for healthcare applications. A key objective is to develop a predictive engine based on these methodologies, ensuring higher accuracy and reliability in forecasting and trend analysis. The approach integrates cutting-edge optimization strategies to improve efficiency and effectiveness in complex systems. CERN will contribute its expertise in Digital Twin technology and advanced methodologies developed in the context of high-energy physics experiments.

In partnership with:

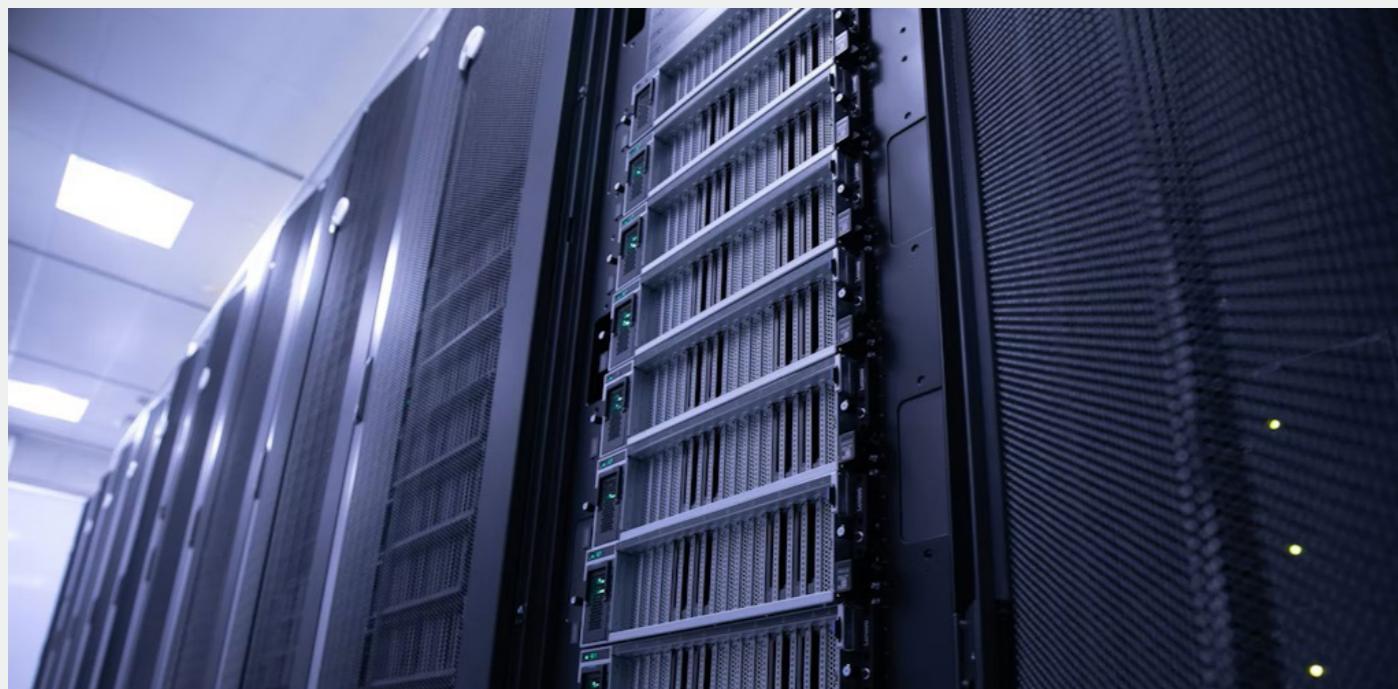
Johnson & Johnson



Applied Multi-Disciplinary AI on High-Performance Computing

Aims to leverage the expertise of both partners in AI optimization and HPC. The expertise in AI optimization workflows suitable for HPC held by CERN, as demonstrated in the use-case of MLPF, will aid scientists at the Simons Foundation Flatiron Institute to leverage AI technologies enabled by modern HPC resources. CERN will study and understand the scientific challenges and applications of scientists at the Simons Foundation Flatiron Institute, and provide guidance and support for optimising suitable AI workflows using HPC platforms. The Flatiron Institute will provide CERN with access to HPC resources, both for the purpose of enabling continued MLPF optimization and for enabling optimization of Simons Foundation use-cases.

In partnership with:



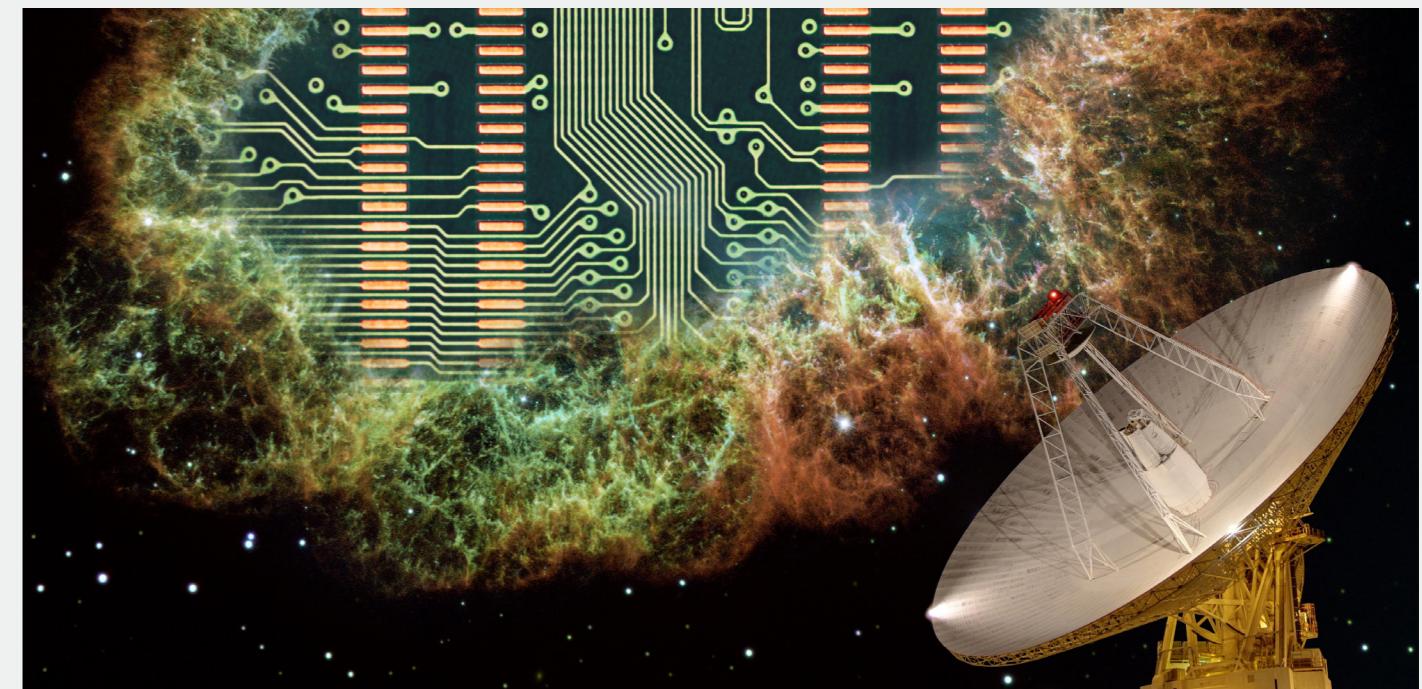
ODISSEE: Online Data Intensive Solutions for Science in the Exabytes Era

This is an EU funded project that federates efforts from 3 pan-European ESFRI infrastructures (HL-LHC, SKAO and SLICESRI) in physical sciences, Big Data, and in the computing continuum supporting flagship instruments that will maintain and strengthen European leadership in high-energy physics and astronomy. The main goal is to enable key science projects, with the search for Dark Matter serving as a pilot program, combining the complementary capabilities of these three unique research infrastructures. ODISSEE will deliver evolutionary and revolutionary hardware and software platforms to address the corresponding digital challenges in a highly competitive international context.

Consortium: Observatoire de Paris, CNRS, CERN, SKAO, ASTRON, INRIA, BSC, Simula, SURF, GENCI, EAS, SiPearl, NextSilicon, NEOVIA, EPFL, CSCS.



This project has received funding from the European Union's Horizon Europe research and innovation program under grant agreement №101188332.



Strategic Partnerships Incubator



CERN openlab as a Strategic Partnership Incubator

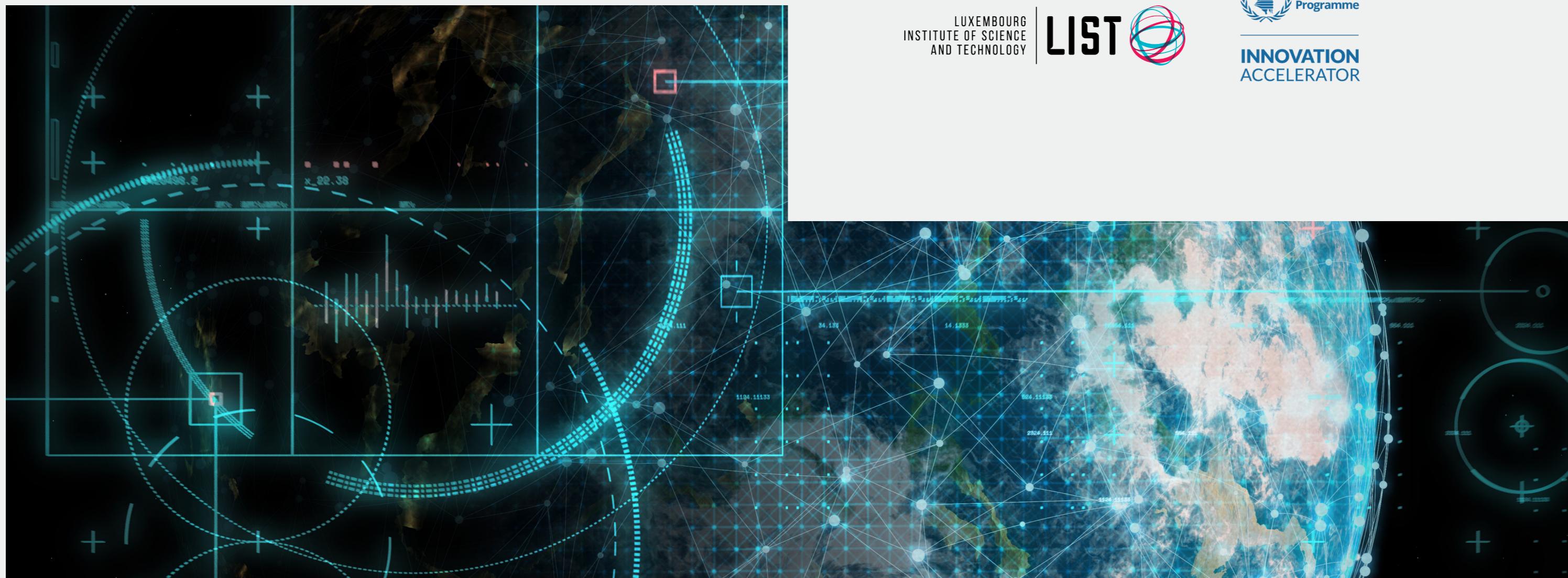
CERN openlab has established a strategic partnerships incubator model by fostering collaborations that leverage cutting-edge computing infrastructures to address scientific and technological challenges. By partnering with leading technology companies, research institutes, and universities, CERN openlab accelerates the development of innovative solutions in areas such as high-performance computing, artificial intelligence, and data storage. Acting as an incubator for these initiatives, CERN openlab paves the way for longer-term partnerships, driving advancements that benefit both the scientific community and industry at large.

Strategic Partnership on Artificial Intelligence

This collaboration between WFP, CERN and LIST offers an unique opportunity to bridge the gap between cutting-edge research and real-world application. By combining WFP's extensive experience in humanitarian operations with CERN's expertise in AI and LIST's capabilities in EO, this project aims to deliver tangible improvements for people facing food insecurity. The strategic partnership on Artificial Intelligence between CERN, WFP, LIST and the Government of Luxembourg is the most recent example of CERN openlab identifying new partnerships and leveraging its implementation model to act as an initial incubator for longer-term collaborations.

In partnership with:

LUXEMBOURG
INSTITUTE OF SCIENCE
AND TECHNOLOGY



Training & Education

In order to prepare for the future of scientific computing, it is vital to ensure that the computing specialists of tomorrow have the right skills to take full advantage of new, innovative technologies. Through projects, lectures and workshops, CERN openlab equips the future computer scientists with indispensable knowledge that inspires scientific advancement and fuels innovation. As a part of the education and training programme, CERN openlab runs various initiatives that support participation of young scientists and other research organisations.



CERN openlab Summer Student Programme

The CERN openlab summer student programme provides undergraduate and master's level students with an opportunity to work on one of the R&D projects for nine weeks under experts' supervision. In addition, the public is granted access to CERN openlab lectures, which cover a broad spectrum of computing subjects, ranging from AI to exascale computing and quantum technologies. To foster community growth, CERN openlab provides ongoing specialised technical training to members of the scientific community.

Collaboration with ideas4HPC

CERN openlab joined forces with ideas4HPC to sponsor a CERN openlab Summer Student, promoting the participation of women in high-performance computing research. ideas4HPC aims to promote inclusivity and diversity in HPC by creating targeted scholarships to support participation in top conferences or running training sessions for mentors. With our summer student programme, we are committed to improving girls' participation in ICT and will continue championing for more women's involvement in ICT.

Lectures, Workshops & Hackathons

Besides the CERN openlab lectures and the summer student programme, CERN openlab projects also contributed to some hackathons. Organised by the CMS collaboration, and with input from CERN openlab researchers working on the project with Micron Technologies, three hackathons were hosted throughout the year at the CERN IdeaSquare, where work on the CXL implementation of the data acquisition software was accelerated and spotlighted.



Technical Workshop

CERN openlab holds an yearly Technical Workshop where members of CERN openlab engage with industry members and the ICT community to showcase the work being done, review of the R&D projects carried out during the past year and discuss future plans. This event features technical talks, poster sessions and technology tracks dedicated to our industrial partners with invited speakers.

In 2024, the Annual CERN openlab Technical Workshop was held at CERN Science Gateway and was attended by more than 130 in-person participants, drawing in experts from multiple companies and research institutes, demonstrating the critical role of public-private partnerships in propelling scientific research.

Participants engaged in rich discussions on the ongoing projects at CERN openlab, delving into the challenges and opportunities at the intersection of technology and science. As well as discussing ongoing projects, the workshop provided an excellent opportunity for considering emerging challenges and identifying opportunities for mutually beneficial collaboration.

The 2024 CERN openlab Technical Workshop presentations encompassed the following topics: Heterogeneous Computing Platforms and Infrastructures, AI applications and HPC Convergence, Applications for Society and Environment, Cloud and Storage, Quantum Computing. Presentations from industry partners were also given.

Maria Girone, the head of CERN openlab, unveiled the ambitious goals of Phase VIII in her presentation. She underlined CERN openlab's unique position as a bridge between industry and scientific research: "Our initiative serves as a unique model for collaboration between science and industry. With Phase VIII, we are committed to pioneering research in emerging technologies and sustainable infrastructures, setting a fast pace for advancements in scientific computing."

As well as discussing ongoing projects, the workshop provided an excellent opportunity for considering emerging challenges and identifying opportunities for mutually beneficial collaboration. This workshop was a venue for presentations and an event to foster future collaborations, showcasing the importance of co-development between industry and the research environment.



Communication & Outreach

CERN openlab serves as a hub of knowledge, dedicated to sharing insights through outreach initiatives and educational programs. In addition to showcasing our technical advancements to diverse stakeholders, we are committed to nurturing the next generation of ICT specialists. By fostering collaboration, CERN openlab enables its members to develop and promote a shared vision for the future of scientific computing. This vision reaches a broad audience, including partner organizations, policymakers, journalists, and the general public. Together, we strive to shape the future of scientific computing to benefit both research and society as a whole.

In 2024, the CERN openlab annual report was revived, a vital tool for communication and outreach, serving as a comprehensive showcase of the collaboration's achievements, initiatives, and future aspirations. It provides a clear and accessible overview of the cutting-edge R&D projects undertaken with our industry partners, highlighting their impact on scientific computing and their contributions to CERN's mission.

This year, we have significantly enhanced our social media presence with the launch of the CERN openlab LinkedIn account, which quickly garnered over 4,000 followers in less than six months. Through LinkedIn, we have shared multiple articles highlighting our projects, achievements, and opportunities, broadening our reach and fostering connections with stakeholders. This increased visibility strengthens our ability to promote CERN openlab's vision and impact, while also inspiring collaboration and innovation in the ICT community.

In the coming year, we will be developing even more exciting communication actions for digital communication with our partners. Through dynamic blog posts, engaging news articles, and captivating social media posts, we are committed to engaging more with our partners to showcase the incredible work we are developing and its importance.

News from the Lab with CERN Alumni

News from the Lab, organised by the CERN alumni team, was designed to showcase the amazing work carried out at the Lab, CERN colleagues are invited to share their work with CERN alumni to help them reconnect with the Organization and its collaborations and keep up to date with the latest news from CERN. In turn, this enables alumni to become CERN ambassadors in their own networks. CERN openlab participated in this event last year and showcased its news, goals and vision for the recently launched Phase VIII. A big thank you to the CERN alumni team for the invite and collaboration.

Head of CERN openlab vice-chairs User Forum Coordination Group aiming to strengthen ties between EuroHPC Users

Post date: 29 Oct 2024

HPC and AI at CERN
Maria Girone
Head of CERN openlab
CERN IT HPC Strategy Coordinator

SUMMER STUDENT PROGRAMME

All the latest from CERN openlab

Project using the BioDynaMo simulation platform among Physics World's Top 10 breakthroughs of 2024: A testament to CERN openlab's vision

SPECTRUM project members gather at CERN to shape Europe's exascale future

Accelerating computing for high-energy physics and beyond: insights from CERN openlab researchers attending supercomputing 2024 in Atlanta

CERN openlab mission

1 | Establishing strategic industry collaborations

2 | Fuelling technological innovation

3 | Exposing technology to researchers

4 | Nurturing knowledge and growth in young STEM researchers

Cerabyte Joins CERN openlab

Cerabyte, the developer of ceramic-based data storage solutions, has announced a strategic partnership with CERN openlab to explore opportunities for AI model training and data storage.

Customer References >

CERN openlab uses OCI for AI model training in its research

CERN openlab tests AI to support physics research scenarios, including building custom large language models with OCI Data Science Service.

Share: [Facebook](#) [X](#) [LinkedIn](#) [Email](#)

"With Oracle, we can build and reuse AI models in a more environmentally friendly way. OCI Data Science makes it easier to share models for projects with multiple detectors and challenges as we experiment with even larger models."

Renato Cardoso, Data Scientist, CERN openlab

Publications & Presentations

T. James, E. Meschi, G. Lazzari Miotto, Real-time Level-1 Trigger Data Scouting at CMS using CXL Memory Lake, Conference on Computing in High Energy and Nuclear Physics, Krakow, Warsaw, Oct 19-25th 2024

T. James, G. Lazzari Miotto, G. Paulino, CXL Memory Management for the CMS L1 Scouting System and Beyond, (Summer Student Report), Dec 2024, <https://zenodo.org/records/14629426>

E. Wulff, J.P. Garcia Amboage, M. Aach, et al., Distributed hybrid quantum-classical performance prediction for hyperparameter optimization. *Quantum Mach. Intell.* 6, 59, Sept 2024, <https://doi.org/10.1007/s42484-024-00198-5>

J. Pata, E. Wulff, F. Mokhtar, et al. Improved particle-flow event reconstruction with scalable neural networks for current and future particle detectors. *Commun Phys* 7, 124, Apr 2024, <https://doi.org/10.1038/s42005-024-01599-5>

J. Pata, E. Wulff, F. Mokhtar, et al. Improved particle-flow event reconstruction with scalable neural networks for current and future particle detectors. *Commun Phys* 7, 124, Apr 2024, <https://doi.org/10.1038/s42005-024-01599-5>

T. Duswald, L. Breitwieser, et al. Calibration of stochastic, agent-based neuron growth models with approximate Bayesian computation. *Journal of Mathematical Biology* 89, 50, Oct 2024, <https://doi.org/10.1007/s00285-024-02144-2>
M. Bunino, K. Tsollaki, et al., itwinai (June). Presented at ISC, Hamburg, 2024.

M. Bunino, K. Tsollaki, et al., interTwin & itwinai (30 April). Presented at EUCAIFCon24, Amsterdam, 2024.

L. Tian, N. Akram, Seismic imaging with remote sensing for energy applications (28 May 2024). Presented at the Third CoE RAISE All-Hands meeting, Barcelona, Spain, 2024.

K. De Grave, Defect-free metal additive manufacturing (28 May 2024). Presented at the Third CoE RAISE All-Hands meeting, Barcelona, Spain, 2024.

E. M. Sumner, Sound Engineering (28 May 2024). Presented at the Third CoE RAISE All-Hands meeting, Barcelona, Spain, 2024.

A. Patil, Data Analytics for Industrial Control Systems (26 March). Presented at CERN openlab Technical Workshop, Geneva, 2024.

D. Southwick, M. Girone, Data Processing Needs and Trends in High Energy Physics, presented at the EGI2024 conference.

T. James, E. Meschi, G. Lazzari Miotto, Real-time Data Processing for CMS Level-1 Trigger using CXL Memory Lake Architecture (26 March). Presented at CERN Openlab Workshop, Geneva, 2024. URL: <https://indico.cern.ch/event/1356148/contributions/5818261/>

E. Wulff, M. Girone, D. Southwick, J.P. Amboage, Center of Excellence for Research on AI- and Simulation-based Engineering at Exascale (CoE RAISE). Poster presented at the 2024 CERN openlab Technical Workshop (26 March 2024), Geneva, Switzerland 2024.

E. Wulff, M. Girone, J.P. Amboage, J. Pata, Distributed Training and HPO at Centre of Excellence on AI and Simulation-Based Engineering at Exascale (CoE RAISE). Presented at the 2024 CERN openlab Technical Workshop (26 March 2024), Geneva, Switzerland 2024.

E. Wulff, Hyperparameter Optimization for Deep Learning using High Performance Computing (9 July 2024). Presented at the CERN openlab Summer Student Lecture Programme, Geneva, 2024.

E. Wulff, Maria Girone, WP4: Data-Driven Use-Cases at Exascale (28 May 2024). Presented at the Third CoE RAISE All-Hands meeting, Barcelona, Spain, 2024.

E. Wulff, M. Girone, J. Pata, Scaling Laws for Machine-Learned Reconstruction (4 June 2024), Poster presented at the PASC24 conference, Zurich, Switzerland 2024.

M. Bunino, A. Zochbauer, et al., interTwin – an Interdisciplinary Digital Twin Engine for Science (26 March). Presented at CERN Openlab Workshop, Geneva, 2024. URL: <https://indico.cern.ch/event/1356148/contributions/5799927/>

J. Santos, L. Atzori, et al., Heterogeneous Architectures Testbed@CERN and E4 project (26 March). Presented at CERN Openlab Workshop, Geneva, 2024. URL: <https://indico.cern.ch/event/1356148/contributions/5799880/>

D. Southwick, L. Atzori, et al., Benchmarking Heterogeneous Architectures with HEPscore (26 March). Presented at CERN Openlab Workshop, Geneva, 2024. URL: <https://indico.cern.ch/event/1356148/contributions/5799887/>

T. James, L. Atzori, et al., A technical overview of industry-science R&D projects for the High Luminosity LHC under CERN openlab (22 October). Presented at CHEP, Krakow, 2024. URL: <https://indico.cern.ch/event/1338689/contributions/6011145/>

C. Pflaum, S. Hellmold, Long-term high-volume data storage in ceramic (14 March). Presented at ACAT, New York, 2024. URL: <https://indico.cern.ch/event/1330797/contributions/5776099/>

E. Wulff, D. Southwick, M. Girone, A. Lektauers, Event Reconstruction and Classification at the HL-HLC (28 May 2024). Presented at the Third CoE RAISE All-Hands meeting, Barcelona, Spain, 2024.

T. James, L. Atzori, et al., A technical overview of industry-science R&D projects for the High Luminosity LHC under CERN openlab (22 October). Presented at CHEP, Krakow, 2024. URL: <https://indico.cern.ch/event/1338689/contributions/6011145/>

C. Rieger, Operational Framework for a Quantum Database (8.4.24), Presentation at TUM Raitenhaslach Seminar, Raitenhaslach, 2024

C. Rieger, Operational Framework for a Quantum Database (17.7.24), Poster presentation at US QIS School Oak Ridge National Lab, Oak Ridge, 2024

C. Rieger, Operational Framework for a Quantum Database (17.9.24), Poster presentation at IQOQI Innsbruck conference, Innsbruck, 2024

C. Rieger, Operational Framework for a Quantum Database (21.1.25), Poster presentation at QT4HEP conference, Geneva, 2025

Rieger, Carla, et al. "Operational Framework for a Quantum Database." arXiv preprint arXiv:2405.14947 (2024).

With thanks to

All partners who have collaborated with CERN openlab activities and everyone who has contributed to the content and production of this document.

Editors

Maria Girone
(Head of CERN openlab)

Mariana Velho
(CERN openlab communication, education and outreach manager)

Graphic Design & Layout

Mariana Velho
(CERN openlab communication, education and outreach manager)

All images designed by Freepik, except:

Page 2, 3, 4, 5, 6, 13, 41, 42, 49, 50, 51, 53, 53, 54: CERN, CERN openlab members, CERN openlab communication office

Page 29: Cerabyte Ceramic-on-Glass Media (Image credit: Cerabyte)

Page 31: A pulsar (pink) can be seen at the center of the galaxy Messier 82 in this multi-wavelength portrait. The pulsar was discovered by NASA's NuSTAR which detected the pulsar's X-ray emission. (Image credit: NASA/JPL-Caltech)

Page 43: The new supercomputer operated by the Flatiron Institute, Simons Foundation (Image credit: Simons Foundation)

Get in touch

 openlab.cern

 openlab-communications@cern.ch

 linkedin.com/showcase/cernopenlab



[LinkedIn](#)



[Website](#)

ISBN

978-92-9083-682-7 (Digital)
978-92-9083-681-0 (Printed)

